Strange bedfellows? Big business meets small farmer

Wytse Vellema

2015

ISBN 978-9-0598984-9-3

Wytse Vellema

Strange Bedfellows?
Big business meets small farmer

UNIVERSITEIT GENT

FACULTEIT BIO-INGENIEURSWETENSCHAPPEN

Foar myn pake, Durk Vellema
(7 februari 1928 - 16 februari 2013)

Cover illustration:
"*The unveiling*"
pencil and gouache on bookcover
PeDeJee
https://pedejee.wordpress.com

Promotor:	Prof. dr. ir. Marijke D'Haese Department of Agricultural Economics Ghent University
Dean:	Prof. dr. ir. Marc Van Meirvenne
Rector:	Prof. dr. Anne de Paepe

Strange bedfellows?

Big business meets small farmer

Thesis submitted in fulfillment of the requirements for the degree of
Doctor (PhD) in Applied Biological Sciences

This thesis will be defended in public on
Friday 18 December 2015 at 16.00

by

Wytse Vellema

Preface

It all started with a big bang, approximately 14 billion years ago. While continents drifted, life started to form, slowly. Some of the first evidence of life on earth is dated between 3.5 to 4.1 billion years ago.[1] Nevertheless, it would still be a long time before humans first came along.

Humans split off from the apes around 7 million years ago, first walked upright about 4 million years ago, and started using primitive tools 2.5 million years ago. Our direct ancestor, homo erectus, came into being around 1.7 million years ago and started spreading out of Africa around 1 million years ago. Sometime after that - nobody knows when exactly - homo sapiens evolved, and spread all over the world.[2]

For more than 99% of the time that modern humans exist, we were living in small bands of hunter-gatherers, semi-sedentary or sedentary in regions with abundant natural resources, but usually migratory. This all changed with the invention of farming, approximately 10,000 years ago. Farming allowed humans to sustain dense populations in small areas and develop societies.

In these initial societies the family was the cornerstone. Production took place within and for the family. However, as farming became more productive societies became larger. Larger societies needed rulers, which needed to be fed from the surplus production of others. In addition, farming required productive assets, and hence protection of property rights. While migratory hunter-gatherers owned only as much as they could carry, the sedentary life-style made possible by the invention of farming allowed for capital accumulation.[3]

It is therefore not surprising that Adam Smith describes this development in the first book ever written on economics (1776).[4] He distinguished four evolutionary stages in the history of civilization: the hunting and pastoral periods of pre-feudal, nomadic cultures, then the farming stage, and finally the commercial era. Smith sees the development of civil society as protection of property rights.

While Smith's view is naturally coloured by the lens of his period and incomplete when it comes to the history of human development, it does bring home an important point. Without farming there is no property, and hence no economics.

The success of farming made human and economic development possible. Given the rate at which the world population is growing, we will rely more than ever before on increases in farm productivity to feed the world. It is only through the exchange of farming knowledge that it is possible to achieve the productivity required for the world to sustain such a large human population. Agricultural development is human development.

Notes

1 The dates for the origin of life are taken from two scientific publications: Bell, E.A., Boehnke, P., Harrison, T.M., Mao, W.L., 2015. Potentially biogenic carbon preserved in a 4.1 billion-year-old zircon. Proceedings of the National Academy of Sciences 112, 14518-14521 and Schopf, J.W., Kudryavtsev, A.B., Agresti, D.G., Wdowiak, T.J., Czaja, A.D., 2002. Laser-Raman imagery of Earth's earliest fossils. Nature 416, 73-76.

2 Boyle, K., Boyle, K.V., Bar-Yosef, O., Stringer, C., 2007. Rethinking the Human Revolution: New Behavioural and Biological Perspectives on the Origin and Dispersal of Modern Humans. McDonald Institute for Archaeological Research.

3 The section on early human societies is based on Diamond, J., 2005. Guns, germs, and steel. Vintage, London, UK..

4 Smith, A., 1776. An inquiry into the nature and causes of the wealth of nations. University of Chicago Press, Chicago, USA.

Members of the jury

Prof. dr. ir. P. Boeckx (Chairperson)
Department of Applied Analytical and Physical Chemistry, Ghent University

Prof. dr. C. Lachat (Secretary)
Department of Food Safety and Food Quality, Ghent University

Prof. dr. ir. M. D'Haese (Promoter)
Department of Agricultural Economics, Ghent University

Dr. D. Boselie (Reading committee)
Senior Manager Learning and Innovation, Sustainable Trade Initiative

Prof. dr. C. P. J. Burger (Reading committee)
Department of Development Economics, Wageningen University

Prof. dr. N. Poole (Reading committee)
Centre for Development, Environment and Policy, University of London

Prof. dr. K. Schoors (Reading committee)
Department of General Economics, Ghent University

Summary

Big business has long been able to avoid dealing directly with small farmers. The high fixed costs of transactions makes dealing with many small farms far costlier than dealing with a few larger farms, leading to a dual system of capitalized large farms delivering to heavily regulated modern supply chains on the one hand, and small farms depending primarily on family labour delivering to local markets on the other. This situation is changing. A rapidly growing and urbanising world population requires agricultural production to increase and become integrated into modern supply chains. Taking more land into use is no longer an option: agricultural activity already covers every nook and cranny of the arable part of the earth. Therefore, increasing output is only possible by increasing productivity on existing agricultural land. The land where potential productivity gains are largest is in the hands of small farmers in low and middle income countries. Most of these farmers do not have the resources or access to credit to close the productivity gap on their own; linking with modern supply chains might prove an opportunity to overcome these limitations and consequently increase their income. The question is how.

The transaction linking small-scale farmers to the next level in the supply chain is called an inclusion mechanism. In the first part of the thesis, the focus was on inclusive business models, a broad subset of inclusion mechanisms. Ten case studies from Burundi, Ethiopia, Kenya, Mozambique, and South Africa with governance structures ranging from almost spot markets to complete vertical integration were analysed using transaction cost theory. In chapter 2, Williamson's transaction characteristics and Ménard's dimensions of hybrids were used to analyse the extent to which existing theory can explain the type and specific form taken by inclusive business models. Results show that although transaction cost theory is a useful starting point for analysing inclusive business models, it is not yet able to perfectly explain the specific form they take, their 'personality'. Financial constraints of participating small-scale farmers was identified as one of the factors which might help to explain this personality.

How exactly financial constraints of farmers affected the personality of inclusive business models was analysed in chapter 3. More specifically, the effect of the combination of firm investment and appropriability hazard on the level and type of safeguards contained within the governance structure was investigated. All studied business models were operating in the context of a weak institutional environment, meaning the local justice system could not be relied upon to cost-effectively mediate conflicts between firm and farmers. In combination with the financial constraints of farmers, this meant that firms had to make the largest share of the investment in the model and find some other way to control

appropriability hazard – the risk of opportunistic behaviour by farmers. Results show that the level of firm investment and the level of safeguards go hand in hand: more investment means more safeguards. However, when appropriability hazard is high, investment is not made in the first place, and a governance structure is adopted which is closer to a take-it-or-leave-it market transaction.

The second part of the thesis focused on measurement of the farm-level impact of inclusion mechanisms. Since small farmers are often poor, the public sector has an interest in finding out the poverty alleviation potential of inclusion mechanisms, to determine if it is worthwhile to invest in them. Unfortunately, measuring this impact is far from straightforward. It requires careful thinking about how and what to measure. In chapter 4 a common measure of impact, the Household Dietary Diversity Score (HDDS), was analysed using household-level data from over 1,000 small farmers on both sides of the border of Colombia and Ecuador. Rasch analysis was used to determine to what extent the indicator was validly measuring food security, one of the outcome targets set in the Millennium Development Goals. Results show that for the sample as a whole, the indicator was not valid. Further analysis showed that its performance could be improved by taking local dietary patterns into account and discarding groups which for particular sub-populations showed abnormal behaviour. Even then, the extent to which the indicator was actually measuring food security was questionable.

Although indicators are a popular method to assess the impact of inclusion mechanisms, there is a risk that such an approach misses broader behavioural impacts. That such broader impacts exist was shown in chapter 5, in which the impact of coffee certification and its corresponding access to specialty coffee markets on the behaviour of small-scale farmers close to the poverty line in the south of Colombia was investigated. The analysis was based on the concept of livelihoods, which holistically considers the set of activities individuals undertake to make a living. Results show that certified farmers obtain a larger share of their income from coffee, because they obtain higher prices and yields and dedicate a larger share of their land to coffee. However, their overall income increases by less than their coffee income. These results might indicate that adhering to certification standards takes so much time and effort that it forces farmers to reduce their participation in other income-generating activities. These broader impacts should be taken into account when assessing the impact of certification.

Considering and analysing inclusion mechanisms as a group, rather than as separate mechanisms, allowed focusing on the way their governance structures solve the farmer inclusion problem. It was shown in chapter 2 that inclusion mechanisms can be analysed using transaction cost theory, although some particular characteristics of the big-business small-farmer transaction they govern - such as the recognition of the importance of financial constraints – require further theory building. In chapter 3, a first step was taken in this direction by

delving deeper into how these financial constraints affect governance structures, and which factors mitigate or strengthen their influence. These chapters constitute the first academic research on the governance structure of inclusive business models.

Two contributions were made to the measurement of farm-level impact. A novel methodology was used in chapter 4 to analyse the relationship between the HDDS food security indicator and dietary patterns of farm households in Colombia and Ecuador. This methodology, Rasch analysis, allows verifying the contribution of each component of the indicator to the underlying latent variable, food security, providing far more insight than traditional methods like Cronbach's alpha, which only shows the overall fit. Moreover, it was the first time the HDDS was tested for validity. In chapter 5, the importance of considering farm-level behaviour was demonstrated. Most research on coffee certification focuses on outcomes which are directly influenced by the certification, such as prices and production. An increase in coffee income is often taken as evidence that the certification improves farmer well-being. However, our research shows its impact has effects beyond those on coffee income and demonstrates how a recently developed econometric technique may be used to measure these.

These findings contain several implications for policy makers. First, that the imbalance in power between firm and farmer, which is often found in inclusion mechanisms, is a direct response to the weak institutional environment. As long as firms cannot rely on public enforcement, the power imbalance might be the price that needs to be paid in order for the firm to make any investment at all. Second, that the ability of the government to overcome this power imbalance – or encourage the firm to include farmers which are poorer than those farmers it prefers to trade with – might be limited. When risks are high, farmers have to be able to demonstrate a credible commitment to the agreement before firms are willing to invest. Such a commitment is not credible when it is made with donated funds. Third, whenever policy makers face a choice between different inclusion mechanisms, the way the alternatives are evaluated requires careful consideration. A choice needs to be made regarding which outcomes to measure and how to measure these outcomes. If indicators are used, the should be adequately verified. In addition, a too narrow focus on a single indicator or outcome area might lead to wrong conclusions.

Table of contents

Preface 5

Members of the jury. 7

Summary 9

Chapter 1
General introduction 17

Chapter 2
Investigating inclusive business models with transaction cost theory 35

Chapter 3
How firm investment and appropriability hazard shape governance structure 59

Chapter 4
Verifying validity of the Household Dietary Diversity Score. 87

Chapter 5
The effect of specialty coffee certification on household livelihood strategies. 125

Chapter 6
General conclusion. 161

Samenvatting voor de leek 181

Acknowledgements 185

List of abbreviations. 191

Curriculum vitae 193

List of figures

Figure 1.1. Stylized depiction of the current organisation of the farm-level impact literature 23

Figure 1.2. Conceptual framework. 24

Figure 1.3. Inclusive Business Models, Inclusion Mechanisms, and the scope of this thesis 27

Figure 4.1. Structure of the results section 97

Figure 4.2. Item response functions (IRFs) Colombia 99

Figure 4.3. Item characteristics curve (ICC) of food group 5 (meat). . . 100

Figure 4.4. Item response functions Ecuador, Kichwa households 102

Figure 4.5. Item response functions Ecuador, migrant households . . . 103

Figure 4.6. Item characteristics curve of food group 7 (fish) for migrant households 104

Figure 4.7. Seasonality in food insecurity for Colombian households . . 109

Figure 4.8. Seasonality in food insecurity for Kichwa households 110

Figure 4.9. Seasonality in food insecurity for migrant households. . . . 110

Figure 5.1. Location of certified and non-certified farms 131

Figure 5.2. Study area: the Department of Nariño, Colombia. 133

List of tables

Table 1.1. Control of activities and asset ownership per governance arrangement. 26

Table 2.1. Control of activities and asset ownership per governance arrangement. 47

Table 3.1. Case studies 71

Table 3.2. Firm investment 73

Table 3.3. Asset specificity and appropriability hazard 74

Table 3.4. Safeguards 76

Table 4.1. Descriptive statistics 93

Table 4.2. Food group consumption by Colombian households. 98

Table 4.3. Food group consumption by Ecuadorian households 101

Table 4.4. Cronbach's alpha statistics for selected Household Dietary Diversity Score groups 105

Table 4.5. Reason for removal and difficulty ranking of food groups for refined indicators 107

Table 4.6. HDDS used in Colombia 112

Table 4.7. HDDS used in Ecuador 113

Table 4.8. 2PL model Colombia including eggs 114

Table 4.9. 2 PL model Kichwa including roots/tubers and fish 114

Table 4.10. 2 PL model migrants including meat and legumes 115

Table 4.11. In- and outfit statistics Colombia (final model) 116

Table 4.12. In- and outfit statistics Kichwa (final model) 116

Table 4.13. In- and outfit statistics migrants (final model) 116

Table 4.14. 1PL model with interaction between item and household groups (DIF) 117

Table 4.15. ANOVA test for DIF 117
Table 4.16. 1PL versus 2PL Colombia 118
Table 4.17. 1PL versus 2PL Kichwa 118
Table 4.18. 1PL versus 2PL migrants. 118
Table 5.1. Certified sales versus certified production 130
Table 5.2. Farm certification and size by municipality 131
Table 5.3. Income distribution between and within income sources . . 134
Table 5.4. Income by source and quintile 135
Table 5.5. Instrument relevance: certification status 140
Table 5.6. Durbin - Wu - Hausman endogeneity tests 141
Table 5.7. Estimated correlation coefficients of the residuals 142
Table 5.8. Participation in specific income-earning activities 143
Table 5.9. Income obtained from specific income-earning activities . . 145
Table 5.10. Income obtained from specific income-earning activities (instrumented) 146
Table 5.11. Total income effects and income pathways 147
Table 5.12. Assumptions power calculation. 151
Table 5.13. Descriptive statistics 152

Chapter 1

General introduction

Problem statement

The growing and urbanising world population is transforming the food system. By 2050, world population is expected to reach 9.6 billion of which 6.3 billion will live in cities (UN Population Division, 2014). As almost all the land suitable for agriculture is already being used, the agricultural production required to feed the world needs to come from increased productivity on existing agricultural land. Most of this increase has to come from low-income countries, where the yield gap – the difference between potential yield and actual yield – is largest (FAO, 2014).

Producing more food is one thing, getting it to consumers quite another. Urbanisation means that food that was traditionally grown and consumed locally increasingly depends on intermediaries in the food value chain. The resulting anonymity between farmer and consumer requires formalized control systems to guarantee food quality and safety (Ruben et al., 2007). Complexity is further increased by the interest of consumer groups in high-income countries in the social and environmental conditions under which food is grown. Taken together, these requirements lead to an increased need for coordination within value chains (Daviron and Vagneron, 2011; Nadvi, 2008; Ponte and Gibbon, 2005; Swinnen, 2007). This raises the question: how can farmers be included in these value chains and how does the inclusion affect them?

A first step to answering this question is to turn it around: which farmers are currently excluded and why? Most research identifies farm size as the most important factor determining exclusion: small farms are often excluded (Barrett et al., 2012; Reardon et al., 2009). Recent estimates suggest there are around 570 million farms in the world (FAO, 2014). Most of these farms are small and family-operated; in low and lower-middle income countries, more than 95% of farms are smaller than five hectares (FAO, 2014). These farms are an important source of national income, especially in low-income countries. Of the 3.1 billion people living in their rural areas, the vast majority depends directly or indirectly on agriculture to make a living (IFAD, 2010; World Bank, 2007). Most are making a meagre living at best; 60% are considered poor and 35% - over 1 billion people - extremely poor[1] (IFAD, 2010). Not all small farms should be considered excluded. Below a certain size, the farm provides insufficient income to sustain a family or even to contribute substantially to their income. Such farmers would benefit equally from opportunities outside their own farm or even outside of agriculture. However, many small or medium-sized farms do have the scale to contribute meaningfully to agricultural production, evidenced by their current production for local markets. Why are so many of these farmers not producing for modern supply chains, which ostensibly offer higher returns? Accessing and using markets

is costly. The costs of using the market are transaction costs, which include per-unit costs such as transportation fees and fixed costs which do not vary with the size of the transaction (Goetz, 1992; Key et al., 2000). Fixed transaction costs include searching for potential partners, bargaining over the terms of the agreement, and enforcing the agreement once it is in place. These costs depend in part on the institutional environment; enforcement costs are higher when the rule of law is weak (Fafchamps, 2004; North, 1990). Transaction costs also differ between marketing channels. Modern supply chains generally have more requirements on quality, safety, and growing conditions than local markets (Maertens and Swinnen, 2009; Ruben et al., 2007). Meeting these requirements frequently involves investments with a large fixed cost component, effectively increasing the scale of production at which it becomes profitable for farmers to participate. Although for many farmers reaching this scale of production is technically feasible by increasing yields, most have insufficient financial resources or access to credit to make the required investments. This combination of high fixed transaction costs, small landholdings, and financial constraints exclude many small farmers from modern supply chains.

Given the growing world population and the consequent increasing demand for food, the need to include small-scale farmers is becoming ever more pressing. Several food companies have recognised this need and have started to include specific targets for smallholder inclusion in their annual reports.[2] Industry-wide learning platforms have also sprung up in which research organisations and the public and private sector share information on best practices. Two examples of such learning platforms are the Seas of Change Initiative and the Sustainable Trade Initiative[3], which have broad membership from food companies. Governments and multilateral donor organisations are playing an important role in funding such initiatives, which are part of a broader shift in the focus of development spending. Governments are increasingly undertaking development projects together with the private sector. In agriculture, several of these projects take the form of 'inclusive business models'[4], which are discussed in this thesis. This is partly for efficiency reasons: it allows NGOs to focus on farmer training, while private partners provide a stable outlet for the farm products. This way, farmers are only encouraged to grow products for which there is an actual demand. On the other hand, the co-investment structure allows governments to maintain the same level of projects while reducing development spending, a process encouraged by the financial crisis. To maximize the effectiveness of development spending and maintain control over the way the funds are used, donors have embraced impact evaluation, in the process making monitoring and evaluation a standard component of almost all development projects. In doing so, a market was created for knowledge on the way such impacts should be measured.

The market for knowledge on farm-level impact

Impact measurement of specific projects and general development methods has rapidly gained prominence over the last decade. Government and multilateral donor interest in aid effectiveness started in earnest with the formulation of the millennium development goals in 2000, which set out specific targets for reducing poverty and improving education and health that the combined development effort should attain by 2015 (UN, 2000). These goals were followed up with the Sustainable Development Goals, which set targets for the next fifteen years.[5] Having such clear targets created the need to be able to objectively assess the contribution of individual development projects to attaining the overall targets, to direct spending to the most effective projects, and thus increase overall aid effectiveness. This ambition towards increased measurement and accountability was formalized in the Paris Declaration on Aid Effectiveness (2005) and signed by governments, multilateral donors, and civil society organisations. Whereas the Paris Declaration focuses on the impact of development projects, the discussion[6] surrounding the process towards aid effectiveness encouraged academic and research organisations working on the agriculture in the developing world to investigate the impact of programs and methods to alleviate poverty leading to a surge of publications on their impact. In agricultural economics, studies on the impact of contract farming and certifications on the well-being of small-scale farmers in developing countries are particularly frequent.

Contract farming is an agreement between a firm and a farmer which gives the firm exclusive buying rights to the contracted crop. Contracts reduce transaction costs by overcoming market imperfections (Bardhan, 1989; Key and Runsten, 1999). In low-income countries, it is common for contracts to provide farmers with seeds, fertilizer, and other inputs at the start of the growing season, which are repaid with part of the harvest (Prowse, 2012). Such inputs are often not available at local markets, and when they are available, farmers frequently do not have the resources or access to credit required to acquire them. Measuring the farm-level impact of contract farming has been far from straightforward, because of limited comparability between contracts in different crops and countries and containing different requirements and support services. Not surprisingly, reviews of the literature find heterogeneous effects (Prowse, 2012; Wang et al., 2014). Impact measurement is complicated further by differences between studies in outcome variables. Some studies measure the effect on production, while others focus on food security. Finally, many contracts specify production requirements based on third-party labels or certifications, such as global GAP, which makes it hard to disentangle to impact of the contract from the impact of the certification. In addition, the contract details are inconsistently provided, making impacts hard to compare across different studies.

Certification is "the action or process of providing someone or something with an official document attesting to a status or level of achievement" (OED, 2008). In agricultural sourcing, this means that by giving a farmer certification, an organisation guarantees that certain conditions have been met in the production of the product. Conditions may take almost any form, from detailed documentation of production practices, abstaining from using certain chemicals in the production process, or membership to a democratic foundation responsible for spending social premiums. The discussion here is restricted to business-to-consumer (B2C) certification - visible on packages in supermarket shelves. B2C certification was born out of the fair trade and organic movements of the 1960s and 1970s (Daviron and Vagneron, 2011; Raynolds, 2000). The first product to receive fair trade certification was coffee[7], which is still the single most important certified crop (Kolk, 2013). Certified products command a premium in the market, which consumers are willing to pay in exchange for the guarantee that products are produced under certain conditions. Part of the premium is used up in the value chain to pay for certifying organisations and marketing, the remainder is redistributed to farmers. Evidence of farm-level impact of certification is limited to specific cases. Most studies focus on the benefits of certification, comparing price, yield, and crop income between certified and non-certified farmers or between different types of certification (Barham and Weber, 2012; Valkila, 2009; Weber, 2011). This emphasis on benefits rather than costs is to be expected, since benefits are easier to measure. The main cost of adhering to strict certification standards is borne by family labour, which is notoriously hard to accurately quantify. Since rural households in developing countries are known to depend on a variety of income sources, all of which require family labour, looking only at those factors directly affected by certification or contract farming might overestimate impact. Moreover, differences in how outcome indicators are measured and which outcome indicators are measured makes it hard to compare impacts between different studies, especially when these studies concern different types of economic relationships.

Bringing together related strands in the literature

The academic literature on farm-level impact of different types of relationships between firms and farmers such as contract farming, certification, and cooperatives is currently organized into separate strands, with little cross-referencing between them. Moreover, there is a tendency to focus on the effect of these separate mechanisms on farm-level outcomes such as food security, production, and poverty. A stylized depiction of the way the literature is currently organized is shown in Figure 1.1. This is not a criticism per se; the statistical rigour required to allow some degree of causal inference requires a clear and narrow

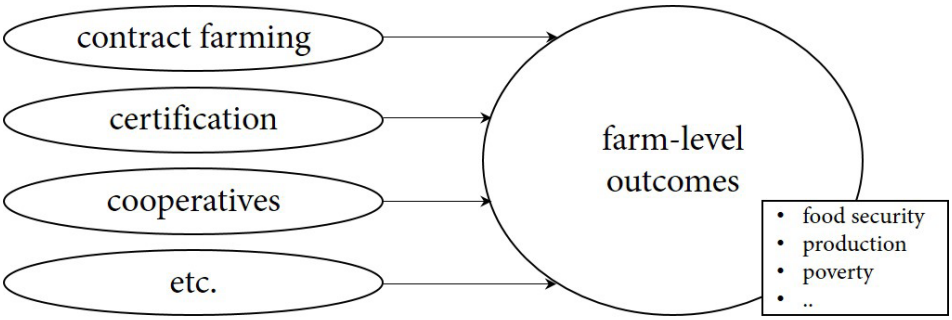

Figure 1.1. Stylized depiction of the current organisation of the farm-level impact literature

focus. Data limitations, often caused by financial and time constraints, exacerbate this tendency. Yet, even if this emphasis on unidirectional and narrow effects is entirely understandable, it is paramount to realize it deviates from reality in at least two important ways.

First, the implied distinction between contract farming, certification, and cooperatives is far from clear-cut. In reality, most forms are diffuse, blending characteristics of one mechanism with those of another, or even completely overlapping. For example, most contract farming involves some process control, which is commonly implemented by requiring farmers to adhere to existing standards such as GlobalGAP (Prowse, 2012). Similarly, Fair Trade International and FLO-Cert only certify coffee producers who are organized in cooperatives[8], making it nearly impossible to differentiate the impact of the certification from the impact of the cooperative. What is problematic is that in practice this overlap is hardly ever made explicit, making the allocation of the study to any of the 'strands' primarily based on the preference of the researcher. Moreover, once the study has been allocated, the probability of it being cited in papers in one of the other strands in the literature drops substantially. Therefore, the current separation between different types of relationship between firms and farmers is not only artificial, it might even be considered an impediment to knowledge building.

Second, considering uniformly applied farm-level outcomes as indicative of true impact is overly simplistic. This tendency is particularly worrying when the outcome is measured with indicators whose validity and reliability has not been verified. Indicators might be particularly sensitive to certain behavioural changes, resulting in invalid comparisons between groups with different preferences. Furthermore, even if the outcome is measured accurately, it might still not be very informative. The finding that food security increased in one instance but not in another is of little value without knowing why the effect was different. And if food security increased, was this because of increased food spending or increased productivity of crops grown for home consumption? In other words, to measure true impact, the interaction between behaviour and outcomes in which household

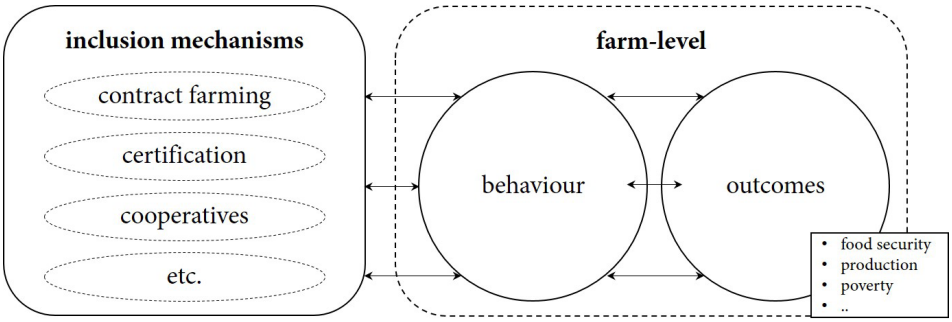

Figure 1.2. Conceptual framework

livelihood choices are likely to play an important role should not be ignored.

In recognition of the limitations associated with the framework in Figure 1.1, in this thesis a more open framework was chosen (Figure 1.2). The different types of relationships between firms and farmers were considered jointly as inclusion mechanisms, defined as organisational structures going beyond arms'-length market transactions governing the economic relationship between a firm and rural households in developing countries, which allows the firm to purchase agricultural products from these households or employ household members in agricultural production or processing.[9] By grouping these mechanisms according to their function – linking small-scale farmers to markets – the analytical focus could be placed on what the different forms have in common and how differences between forms might be explained.

Measurement issues

Accurate measurement of farm-level impact is complex and recent years have seen substantial progress in methodological developments. In agricultural and development research, most of this effort has been concentrated on improved targeting and sampling (Banerjee and Duflo, 2011; Ravallion, 2008). The main objective of sampling is to select study participants such that the sample is an unbiased representation of the study population of interest. Targeting concerns the way beneficiaries of a particular intervention are selected. Adequate sampling and targeting ensures that the impact of the intervention is measured in isolation from other factors affecting the outcome variable.

A topic which has received less attention, at least in the field of agricultural economics, is indicator validity and reliability. Validity refers to "the extent to which an empirical measure adequately reflects the real meaning of the concept under consideration" (Babbie, 1989), while "a scale or test is reliable to the extent that repeat measurements made by it under constant conditions will give the

same result" (Moser and Kalton, 1989). Validity and reliability matter because of the increased use of indicators to measure progress towards the targets of the Millennium Development Goals, which focus on food security, poverty, and health outcomes. Impact measurement is expensive and uses resources which could have been used for operational activities generating the impact. To save resources, rather than collect detailed income or food consumption data to determine poverty and food security status, indicators are used: a limited set of questions that allow rapid assessment. Such indicators are particularly common for poverty and food security. Unfortunately, tests of their validity and reliability are scarce (Desiere et al., 2015; Heady and Ecker, 2013), yet, such tests are an essential first step to verify whether these indicators measure what they were developed to measure.

An intervention, whether it is a development project or the inclusion of a farmer in the value chain, has a myriad of impacts. Although the Millennium Development Goals might encourage aid effectiveness by increasing the emphasis on outcome measurement, at the same time they encourage a narrow focus on a limited set of outcomes. Such a focus is problematic when it prevents a deeper understanding of long-term behavioural impacts of the intervention (Pritchett and Woolcock, 2004). There is a difference between distributing monthly food rations and providing households with the capability to grow the same amount of food that is not picked up when only values on ultimate outcomes like poverty or food security indicators are considered. Thankfully, most researchers interested in the impact of specific inclusion mechanisms are not bound by such constraints. And yet, even in this literature there appears to be a preference for relatively easy-to-observe outcomes like price, production, and marketed surplus (e.g. Barham and Weber (2012), Prowse (2012), Valkila (2009)). This narrow focus is particularly surprising given the now well-established finding that off-farm income sources play an important role for small-scale farmers (Ellis, 1998, 2000) and that rural households rely on a broad set of income-generating activities to sustain their livelihood (Davis et al., 2010; Scoones, 1998).

When it comes to the impact of inclusion mechanisms, such broader impacts deserve to be taken into account. If, perhaps because of time or financial constraints, the decision is made to rely on indicators to evaluate impact, these indicators should first be validated. It should also be clear which specific inclusion mechanism is evaluated to be able to assess how it compares to available alternatives. To define what constitutes such an available alternative, more research is needed on the governance structures of inclusion mechanisms. Such research starts with a clear definition.

Inclusive business models, inclusion mechanisms, and the scope of this thesis

There are numerous organisations working on inclusive business models, many of which choose to use their own definition. An overview of some of the most current definitions as well as the organisation using them is provided in Table 1.1. Although the definitions differ in their wording and sometimes in their emphasis, there are at least four aspects they have in common. First, their focus on the poor, variously defined as small-scale farmers (1), low-income people (2, 8, 9), or those at the base of the pyramid (3, 4, 5). Second, an acknowledgement of the necessary win-win for both the company and those it includes. This is framed in

Table 1.1. Control of activities and asset ownership per governance arrangement

	Definition	Institution
1	Those [business models] which do not leave behind small-scale farmers and in which the voices and needs of those actors in rural areas in developing countries are recognised (Vorley et al., 2008).	FAO
2	Inclusive businesses include low-income people on the demand side as customers, and on the supply side as employees, producers and entrepreneurs serving at various points within the value chain. They build bridges between business and the poor for mutual benefit (UNDP, 2008).	UNDP
3	Inclusive business models expand access to goods, services, and livelihood opportunities for those at the base of the pyramid in commercially viable, scalable ways (Jenkins et al., 2011).	IFC
4	Inclusive business models are a private sector approach to providing goods, services and livelihoods on a commercially viable basis, either at scale or scalable, to people at the base of the pyramid by making them part of the value chain of companies' core business as suppliers, distributors, retailers or customers (Hertveldt et al., 2012).	IFC, G20
5	A profitable core business activity that also tangibly expands opportunities for the people at the base of the economic pyramid (BoP): as producers, suppliers, workers, distributors, consumers – or even as innovators (BIF, 2012).	BIF, DFID
6	Inclusive business models include the poor on the demand side as clients, and on the supply side as distributors, suppliers of goods and services, or employees at various points in the value chain. In so doing, these business models build bridges between business and the poor for mutual benefit (Tewes-Gradl et al., 2013).	Endeva
7	Inclusive business models make a positive contribution to the development of companies, the local population and the environment (Gradl and Knobloch 2010).	Endeva
8	An inclusive business is an entrepreneurial initiative seeking to build bridges between business and low-income populations for the benefit of both (WBCSD and SNV, 2008).	SNV-WBCSD
9	An inclusive business is an economically profitable, environmentally and socially responsible entrepreneurial initiative, which integrates low-income communities in its value chain for the mutual benefit of both the company and the community. It seeks to improve the livelihoods of low-income populations while increasing returns to the company (…) (SNV-WBCSD, 2011).	SNV-WBCSD

terms of mutual benefit (2, 6, 8, 9), commercial viability (3), or profitability (5, 9). Third, the stated interest of the for-profit business partner in the well-being of the poor which are included through the business model, recognising their voices and needs (1), contributing to development (7), and seeking to improve their livelihoods (9). Fourth, the mechanism used to include goes beyond an arms-length market transaction, reflected by terms like making them part of (4), building bridges (2, 6, 8) and integrate (9). The idea is that certain vulnerable populations are excluded from participating in the mainstream market, and that something extra is required to allow them to participate.

In this thesis, inclusive business models are defined as private sector for-profit initiatives which include the poor with the stated intention to improve their well-being through mechanisms going beyond arm's-length market transactions.

At first sight, there appears to be little difference between inclusive business models and inclusion mechanisms. Both include small-scale farmers, require profit for both the firm and the farmer in order to be durable, and go beyond an arms'-length, spot market transaction. This overlap means that in many cases, a model can be an inclusive business model and an inclusion mechanism at the same time (Figure 1.3). However, there is one big difference between the two. Inclusive business models require something which inclusion mechanisms do not: the stated intention to make the included population better off. Without such a stated intention, an inclusion mechanism should not be considered an inclusive business model but a 'commercial and inclusive value chain' (Harper et al., 2015), whose definition explicitly excludes inclusion mechanisms which have increasing the well-being of the people they include as an objective. Another difference between inclusive business models and inclusion mechanisms is that the former may also concern consumers. Such consumer-oriented models fall outside the scope of this thesis.

All cases studied in this thesis concern inclusion mechanisms – as defined in relation to Figure 1.2 – which may also be considered inclusive business models, because of their stated interest in the well-being of the farmers they attempt

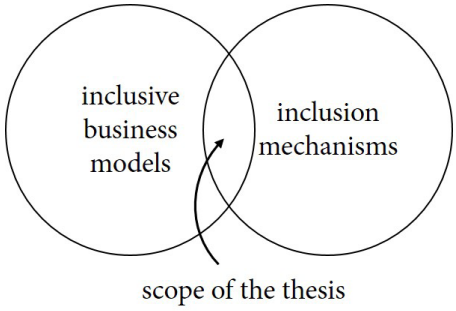

Figure 1.3. Inclusive Business Models, Inclusion Mechanisms, and the scope of this thesis

to include. This is not to say that the well-being of farmers does not matter for other inclusion mechanisms, quite the contrary. When farmers do not expect to benefit, they will not participate in any inclusion mechanism (Barrett et al., 2012). And when expected benefits do not materialize, they are unlikely to be willing to continue the relationship (Gow and Swinnen, 2001). Hence, even those inclusion mechanisms which include small farmers purely out of self-interest may logically be expected to have a positive effect on the well-being of farmers. Therefore, even though the studies in this thesis exclusively concern inclusive business models, which are a subset of inclusion mechanisms, the results are expected to extend to most non-IBM inclusion mechanisms.

Thesis outline

The overall objective of this thesis was to contribute to the understanding of the way small-scale farmers in developing countries are included in globalized food value chains and the measurement of the impact of this inclusion on their well-being. In researching the subject, it became clear why smallholders are currently excluded from these value chains, and that there exist many different mechanisms for smallholder inclusion. A diverse sub-group within these inclusion mechanisms, inclusive business models, were studied in this thesis. What sets inclusive business models apart from other inclusion mechanisms is their stated interest in the well-being of the rural households they include, which makes accurate and complete measurement of their impacts on this well-being all the more relevant. The thesis is divided into two parts and contains four independent chapters: the first part considers the governance structure of inclusive business models, the second part the measurement of farm-level impact.

This thesis is built on two sets of case studies, one set with data at the level of the governance structure, the other with farm-level data. The first data set on governance structures contains ten case studies of inclusive business models in Burundi, Ethiopia, Kenya, Mozambique, and South Africa. In chapter 2, one of these ten case studies was selected for in-depth within-case analysis, which is a recommended approach to generate detailed understanding of complex organisational processes (Eisenhardt, 1989; Hartley, 2004; Yin, 1994). In chapter 3, the other ten case studies were compared using across-case analysis, which is useful to identify differences and similarities across cases (Eisenhardt, 1989). The second data set contained farm-level data from around 1,000 households in Colombia and Ecuador. In chapter 4, data from both countries was used to check the internal validity of a food security indicator which is frequently used to measure the farm-level impact of development interventions. In chapter 5, data from Colombia was used to test the impact of farm certification – an inclusive business model – on the livelihood strategies of small-scale farmers.

In chapter two, the same theory which was used to identify the reason for the frequent exclusion of small-scale farmers from globalized food value chains, transaction cost theory, is applied to three distinct governance structures governing the purchase of sugar cane from small-scale farmers in KwaZulu-Natal, South Africa. The objective of this chapter is to discover the extent to which transaction cost theory, developed mostly for businesses in high-income countries, can explain governance structures in inclusive business models. Specifically, it applies Williamson's (1975) transaction attributes of frequency, uncertainty, and asset specificity and Ménard's (2004) dimensions of monitoring mechanisms, rent allocation, and enforcement mechanisms to explain the specific form taken by these governance structures. Financial constraints of farmers were identified as an important explanatory factor of their form, not taken into account by existing theories on governance structures.

In chapter three, the effect of the financial constraints of farmers - and the consequent onus to invest placed on firms - on the governance structure of a set of ten case studies of inclusive business models in Africa was studied. All cases share a context of a weak institutional environment requiring the firms to rely on safeguards contained in the governance structure to ensure contract compliance. Transaction cost theory was complemented by positive agency theory to investigate which factors determine the risk of the investment, and the extent to which safeguards are able to mitigate these risks.

The second part of the thesis concerns the measurement of farm-level impact. In chapter four, the validity of a frequently used indicator of food security, the Household Dietary Diversity Score (HDDS), was tested. This indicator was developed in 2006 (Swindale and Bilinsky, 2006) and is widely used. However, its validity had never been tested, creating the risk that the indicator might not actually measure what it was developed to measure: household food access, a dimension of food security. Rasch analysis was used to test the indicator's internal validity, i.e. the extent to which the questions of which the indicator is made up appear to be related to the same underlying construct. Results are disappointing, and harbour potentially bad news for donors which have been relying on this indicator to evaluate programs.

In chapter five, the effect of coffee certification on the livelihood of small-scale coffee producers in the south of the Colombian Andes was studied. Certification of small-scale farmers is a classic example of an inclusion mechanism which states to be interested in the well-being of farmers. Rural households are known to rely on many different activities to generate their income. The objective of this chapter is to analyse if certification had effects on the income-generating activities other than coffee production, thus explicitly going beyond directly observable outcomes, such as price and productivity. The relative attractiveness of coffee production with certification is analysed using the economic principle of revealed

preference. Results show certification indeed increases the relative attractiveness of coffee production, but that increases in income from coffee are at the expense of income from other sources.

References

Babbie, E., 1989. Survey research methods, 2nd ed. Wadsworth, Belmont, US.

Banerjee, A.V., Duflo, E., 2011. Poor economics: a radical rethinking of the way to fight global poverty. PublicAffairs, United States.

Bardhan, P., 1989. The economic theory of agrarian institutions. Oxford University Press, New York, US.

Barham, B.L., Weber, J.G., 2012. The Economic Sustainability of Certified Coffee: Recent Evidence from Mexico and Peru. World Development 40, 1269-1279.

Barrett, C.B., Bachke, M.E., Bellemare, M.F., Michelson, H.C., Narayanan, S., Walker, T.F., 2012. Smallholder Participation in Contract Farming: Comparative Evidence from Five Countries. World Development 40, 715-730.

BIF, 2012. What is Inclusive Business? Innovations against poverty. Business Innovation Facility.

Cargill, 2014. Corporate responsibility report.

Daviron, B., Vagneron, I., 2011. From Commoditisation to De-commoditisation … and Back Again: Discussing the Role of Sustainability Standards for Agricultural Products. Development Policy Review 29, 91-113.

Davis, B., Winters, P., Carletto, G., Covarrubias, K., Quiñones, E.J., Zezza, A., Stamoulis, K., Azzarri, C., Digiuseppe, S., 2010. A cross-country comparison of rural income generating activities. World Development 38, 48-63.

Desiere, S., Vellema, W., D'Haese, M., 2015. A validity assessment of the Progress out of Poverty Index (PPI)™. Evaluation and Program Planning 49, 10-18.

Eisenhardt, K.M., 1989. Building theories from case study research. Academy of Management Review 14, 532-550.

Ellis, F., 1998. Household strategies and rural livelihood diversification. The Journal of Development Studies 35, 1-38.

Ellis, F., 2000. Rural livelihoods and diversity in developing countries. Oxford University Press, Oxford.

Fafchamps, M., 2004. Market institutions in Sub-Saharan Africa : theory and evidence / Marcel Fafchamps. MIT Press, Cambridge, Mass. :.

FAO, 2014. The state of food and agriculture 2014. UN Food and Agriculture Organisation.

Goetz, S.J., 1992. A Selectivity Model of Household Food Marketing Behavior in

Sub-Saharan Africa. American Journal of Agricultural Economics 74, 444-452.

Gow, H.R., Swinnen, J.F.M., 2001. Private Enforcement Capital and Contract Enforcement in Transition Economies. American Journal of Agricultural Economics 83, 686-690.

Harper, M., Belt, J., Roy, R., 2015. Commercial and inclusive value chains: Doing good and doing well. Practical action publishing, Bourton on Dunsmore, UK.

Hartley, J., 2004. Case study research, in: Cassell, C., Symon, G. (Eds.), Essential guide to qualitative methods in organisational research. Sage, London, UK, pp. 323-333.

Heady, D., Ecker, O., 2013. Rethinking the measurement of food security: From first principles to best practice. Food Security 5, 327-343.

Hertveldt, S., Geaneotes, A., Masuoka, T., 2012. Policy note on the business environment for inclusive business models. International Finance Corporation, Washington, DC.

IFAD, 2010. Rural poverty report 2011. International Fund for Agricultural Development, Rome, Italy.

Jenkins, B., Ishikawa, E., Geaneotes, A., Baptista, P., Masuoka, T., 2011. Accelerating inclusive business opportunities: Business models that make a difference. International Finance Corporation (IFC), Washington, DC.

Key, N., Runsten, D., 1999. Contract Farming, Smallholders, and Rural Development in Latin America: The Organisation of Agroprocessing Firms and the Scale of Outgrower Production. World Development 27, 381-401.

Key, N., Sadoulet, E., Janvry, A.D., 2000. Transactions Costs and Agricultural Household Supply Response. American Journal of Agricultural Economics 82, 245-259.

Kolk, A., 2013. Mainstreaming sustainable coffee. Sustainable Development 21, 324-337.

Maertens, M., Swinnen, J.F.M., 2009. Trade, Standards, and Poverty: Evidence from Senegal. World Development 37, 161-178.

Moser, C., Kalton, G., 1989. Survey methods in social investigation, 2nd ed. Gower, Aldershot, US.

Moyo, D., 2009. Dead aid: why aid is not working and how there is a better way for Africa. Farrar, Straus and Giroux, New York, US.

Nadvi, K., 2008. Global standards, global governance and the organisation of global value chains. Journal of Economic Geography 8, 323-343.

North, D.C., 1990. Institutions, institutional change, and economic performance. Cambridge University Press, Cambridge, UK.

OED, 2008. Concise Oxford English Dictionary, 11th ed. Oxford University Press.

Paris Declaration on Aid Effectiveness. 2005.

Ponte, S., Gibbon, P., 2005. Quality standards, conventions and the governance of global value chains. Economy and Society 34, 1-31.

Pritchett, L., Woolcock, M., 2004. Solutions when the Solution is the Problem: Arraying the Disarray in Development. World development 32, 191-212.

Prowse, M., 2012. Contract farming in developing countries - A review, in: Peccoud, R. (Ed.), A Savoir. Imprimerie de Montligeon, France.

Ravallion, M., 2008. Evaluation anti-poverty programs, in: Schultz, T.W., Strauss, J. (Eds.), Handbook of Development Economics. Elsevier.

Raynolds, L., 2000. Re-embedding global agriculture: The international organic and fair trade movements. Agriculture and Human Values 17, 297-309.

Reardon, T., Barrett, C.B., Berdegué, J.A., Swinnen, J.F.M., 2009. Agrifood Industry Transformation and Small Farmers in Developing Countries. World Development 37, 1717-1727.

Ruben, R., Van Boekel, M., Van Tilburg, A., Trienekens, J., 2007. Tropical Food Chains. Wageningen Academic Publishers, Wageningen, NL.

Scoones, I., 1998. Sustainable rural livelihoods: A framework for analysis, IDS Working paper. Institute of Development Studies.

SNV-WBCSD, 2011. Inclusive Business: Creating value in Latin America.

Swinnen, J., 2007. Global supply chains, standards, and the poor: how the globalization of food systems and standards affects rural development and poverty. Cromwell Press, Trowbridge, UK.

Tewes-Gradl, C., Peters, A., Vohla, K., Lütjens-Schilling, L., 2013. Inclusive business policies. How governments can engage companies in meeting development goals. Endeva UG.

UN, 2000. United Nations Millennium Declaration, 55/2, General Assembly ed, p. 60(b).

UNDP, 2008. Creating Value for All: Strategies for Doing Business with the Poor. Report of the Growing Inclusive Markets Initiative. United Nations Development Programme, New York.

Unilever, 2014. Strategic report, Annual report and accounts.

Valkila, J., 2009. Fair Trade organic coffee production in Nicaragua — Sustainable development or a poverty trap? Ecological Economics 68, 3018-3025.

Vorley, B., Lundy, M., MacGregor, J., 2008. Business models that are inclusive of small farmers, Global Agro-Industries Forum, New Delhi.

Wang, H.H., Wang, Y., Delgado, M.S., 2014. The Transition to Modern Agriculture: Contract Farming in Developing Economies. American Journal of Agricultural Economics 96, 1257-1271.

WBCSD, SNV, 2008. Inclusive business: Profitable business for sustainable development. World Business Council for Sustainable Development (WBCSD); SNV Netherlands Development Organisation.

Weber, J.G., 2011. How much more do growers receive for Fair Trade-organic coffee? Food Policy 36, 678-685.

World Bank, 2007. World Development Report 2008: Agriculture for Development. World Bank, Washington D.C.

Yin, R.K., 1994. Case study research: Design and methods. Sage, Thousand Oaks, USA.

Notes

1 A person was considered poor if his or her income was less than 2 USD/day and extremely poor when it was less than 1.25 USD/day

2 See for example the annual reports for 2014 of Cargill and Unilever.

3 Websites: www.seasofchange.com and www.idhsustainabletrade.com.

4 "Private sector for-profit initiatives which include the poor with the stated intention to improve their well-being through arms'-length market transactions.

5 http://un.org/sustainabledevelopment/sustainable-development-goals.

6 The fire was fuelled by publication of Moyo's book 'Dead Aid' (2009), in which the author argues that past aid has been not merely ineffective, but even counterproductive.

7 Mexican coffee, to be precise. In 1988 the first fair trade coffee appeared in Dutch supermarkets under the Max Havelaar label.

8 On 31 December 2011, FT USA split from Fair Trade International because it believed non-organized coffee producers should also be eligible for certification. See www.fairtradeusa.org for the official announcement and explanation (15 September 2011).

9 This definition was developed for the purpose of this thesis. It explicitly includes rural households as employees, to allow for changes in agricultural production systems such as those described by Maertens and Swinnen (2009).

Chapter 2

Investigating inclusive business models with transaction cost theory

based on

Vellema, W., D'Haese, M., 2015. *Explaining hybrid "personalities" in smallholder sugar cane sourcing*. British Food Journal 117, 2547-2563.

Abstract

Inclusive business models are a diverse set of governance structures linking small-scale farmers to modern food value chains, often taking the form of public-private partnerships or private sector developments with poverty alleviation or increasing food security as their specific objective. Although these models have been described extensively in case studies conducted by donors and development organisations, there have been few attempts to compare cases and link them to theory. In this chapter, an attempt is made to fill this void. More specifically, to investigate to what extent transaction cost theory can be used to analyse inclusive business models as a single heterogeneous group. Three governance structures used simultaneously by the same company to purchase sugar cane from small-scale growers are described in detail. One of these structures is close to a market arrangement, the other two are hybrids. Williamson's discriminating alignment hypothesis and Ménard's recent work on hybrid models are used to explain the factors driving the choice for a hybrid arrangement and determining their specific form. Results indicate that at least two areas would need to be included to explain the specific form taken by the studied governance structures: production characteristics and financial constraints of the transacting parties. In addition, national and local regulations are shown to have an important effect on organisational form. Direct government involvement in designing the governance structure also had a large influence on the degree of inclusiveness. However, some care is warranted: overly restrictive inclusion criteria might have an adverse effect by endangering the long-term survival of the model.

Introduction

Institutional and regulatory changes are affecting agri-food value chains worldwide. Although some trends and changes might be global in nature (Reardon et al., 1999), the way organisations are affected depends equally on the country-specific or even local institutional environment. In South Africa, there is a history of extensive government intervention in agricultural markets (Kirsten et al., 2009). The end of apartheid marked a strong shift in policy towards re-integrating the formerly excluded black population into the productive process. One of the most important pillars of this policy in the agricultural sector was land reform, combined with extensive opportunities for government to encourage participation of small landholders in agri-food value chains. These recent policy shifts have had a profound influence on the organisation of the agricultural sector.

An understanding of the organisation of any economic activity, including agriculture, requires an explanation of the conditions determining why some activities are performed jointly while others are performed independently. Such an understanding ranges from the division of labour in a needle factory (Smith, 1776) to the boundaries of the firm (Coase, 1937). These boundaries can be studied by considering the transaction as the unit of analysis, in particular the so-called make or buy decision. In his 1975 book, Williamson (1975) distinguishes two main governance structures for transactions: markets and hierarchies. In transactions governed by markets, asset ownership and decision rights are split between two or more autonomous firms, whereas in hierarchies ownership and decision-making are in the hands of a single firm.

Empirical studies on governance structures have shown that there are more than two ways to govern transactions (Blois, 1972). In the economics literature, these other governance structures are jointly named hybrids, a term coined by Rubin (1978). They are defined as 'arrangements in which two or more partners pool strategic decision rights as well as some property rights, while simultaneously keeping distinct ownership over key assets, so that they require specific devices to coordinate their joint activities and arbitrate the allocation of payoffs' (Ménard, 2004) and are considered an intermediate form in the continuum between markets and hierarchies (Williamson, 1985). As such, they are commonly analysed by contrasting them with markets and hierarchies, using the transaction characteristics of asset specificity, uncertainty, and frequency originally developed by Williamson (1991).

However, these 'mainstream' characteristics might be insufficient to explain the specific form taken by hybrids, their 'personality' (Ménard, 2012a). Theoretical and applied research has identified a large variety of hybrid governance structures, with terminology and classification depending on whether the study was conducted within the field of economics, management, or sociology. This

variety with partly overlapping categories has complicated advances on the study of factors driving the choice of hybrid structure. One of the first attempts in this regards was made by Ménard (2004), who used monitoring mechanisms, rent allocation, and enforcement mechanisms as identifying characteristics. The extent to which these dimensions explain the specific form taken by hybrids is one of the main objectives of this paper.

Existing theory on hybrids will be applied to a case study containing two agricultural sourcing arrangements which may be considered inclusive business models, so-called because of their explicit inclusion of the poor or disadvantaged (UNDP, 2008). Inclusive business models are rapidly gaining prominence in development programs, yet have received little consideration in academic thinking. Because these models more often than not take hybrid forms they fall within the scope of transaction cost theory. Yet, their complexity raises questions on the extent to which commonly used elements of the transaction cost lens enable reading and understanding these inclusive business hybrids. All of the arrangements studied in this paper govern the same transaction by the same company - sugar cane sourcing from small-scale growers - but differ in their details. Hence, the case study allows a clear distinction between common and unique elements. The aim of this paper is to understand the 'personality' of each of these structures, and the way in which they have been shaped by South Africa's institutions, in particular regarding industry policies, regulated price setting, and land reform.

Before the description of the transaction and the governance structures used to govern it, the method used to analyse the case and the data collection is explained. After that, the characteristics of the studied transaction is described in terms of Williamson's (1991) discriminating alignment hypothesis, focusing on frequency, uncertainty, and asset specificity. The fourth section contains historical and descriptive information on relevant aspects of the institutional environment. Particular attention is given to historical and current sugar industry-specific regulation, as well as the local implementation of the land reform policy. The studied governance structures are described in more detail in section five, comparing control over activities and ownership of assets by each party. In the sixth section the relationship between transaction characteristics and governance structures is studied, looking first at reasons to go hybrid and second at the personality of the specific forms chosen. In the final sections main findings are discussed and summarized.

Method

The different sourcing arrangements were analysed in-depth using the rationale of discrete structural analysis (Simon, 1978). Williamson (2002) argues in favour of this approach to build understanding of complex microeconomic organisation.

The principal burden of analysis should be based on comparisons of transaction costs between economic institutions. Therefore the analysis will be based on a combination of within-case and cross-case analysis. The case study method is well suited for in-depth study of a phenomenon in its real-life context (Yin, 1994), particularly to research questions requiring detailed understanding of complex social or organisational processes (Hartley, 2004). In this paper we combine within-case with cross-case analysis. The first is helpful for in-depth analysis and description, the latter to analyse similarities and differences across cases (Eisenhardt, 1989). It is therefore a descriptive and comparative approach.

Inclusive business models are a relatively recent but increasingly common phenomenon, which has received little attention in the academic literature. It was therefore not possible to directly apply or test existing theory. Therefore, in this paper a case study approach was chosen, in combination with a broadly scoped research question, namely 'to what extent can existing theory explain the specific form taken by inclusive business models and what are possible areas that require further study?'. Such a broadly scoped question was considered appropriate given the paucity of existing theory (Eisenhardt and Graebner, 2007).

The selected case study was part of a broader set of ten case studies on inclusive business models, described in detail in Sopov et al. (2014). These case studies were selected not to be representative of the population of inclusive business models, but rather to demonstrate the variety within the population. Since the project constituted the first attempt to consider these models as a group, few a priori criteria existed. Therefore, a more explorative approach was chosen. Research and practitioner organisations in sub-Saharan Africa were contacted and asked to provide short descriptions of inclusive business models they were familiar with. Out of these descriptions, ten cases were selected that were (1) as diverse as possible regarding their governance structures and (2) included different products, with different end-markets, in different parts of the continent. The full set of case studies is described in more detail in Chapter 3.

From the broader set of ten case studies in the research project, sugar cane sourcing in South Africa was selected. This case was considered to be particularly relevant to study how transaction cost theory might be applied to the phenomenon of inclusive business models, since it contains multiple governance structures governing the same transaction. Therefore, it allows a differentiation between common and unique elements which is not possible when comparing governance modes between different transactions or institutional settings. An additional motivation to select this specific case was that it is located in South Africa, which provides a well-developed institutional and regulatory environment. This density of the institutional environment generates variation, which allows discovery of the finer interactions with organisational form. Furthermore, due to South Africa's long history of well-developed institutions, information on policies and

regulations was well-documented an easily accessible.

To collect detailed information on the different governance structures and local situation, face-to-face interviews were held on location. Such an interview-based method is particularly useful to generate insights into how and why questions (Pettigrew, 1990). Interviews were complemented with desk research for triangulation (Yin, 1994), especially on relevant laws and regulations. Interview protocols were standardized to facilitate comparisons (Miles and Huberman, 1994). Using pre-developed questions also allowed maintaining a link between existing knowledge and inductive elements of the study (Pettigrew, 1990). Questions were built on the LINK methodology, which was developed to understand inclusive trading relationships between farmers, farmer organisations, and formal markets (Lundy et al., 2012).

Data collection for the cooperative model took place over a two-week period in October 2013 and was initiated via emails explaining the objective of our study. The first interviews were held with the company managers responsible for developing the cooperative model. During these interviews, we learned about the existence of the third-party model. Interviews for this model, which was located in a different area, were conducted in August 2014. Interviews with managers were followed up by interviews with board members of successful and less successful cooperatives, field staff, and contractors. For the third party model, interviews were also held with managers and employees of the management company. After having interviewed the stakeholders, follow-up interviews were held with company management to clarify issues that came up. During writing up, additional contact by email and telephone was used to check details and obtain missing information.

Data

The selected case study was part of a broader set of ten case studies on inclusive business models, described in detail in Sopov et al. (2014). These case studies were selected not to be representative of the population of inclusive business models, but rather to demonstrate the variety within the population. Since the project constituted the first attempt to consider these models as a group, few a priori criteria existed. Therefore, a more explorative approach was chosen. Research and practitioner organisations in sub-Saharan Africa were contacted and asked to provide short descriptions of inclusive business models they were familiar with. Out of these descriptions, ten cases were selected that were (1) as diverse as possible regarding their governance structures and (2) included different products, with different end-markets, in different parts of the continent. These case studies are described in detail in chapter 3.

From the broader set of ten case studies in the research project, sugar cane sourcing in KwaZulu-Natal, South Africa, was selected for an in-depth analysis.

This case was considered to be particularly relevant to study how transaction cost theory might be applied to the phenomenon of inclusive business models, since it contains multiple governance structures governing the same transaction. Therefore, it allows a differentiation between common and unique elements which is not possible when comparing governance modes between different transactions or institutional settings. An additional motivation to select this specific case was that it is located in South Africa, which provides a well-developed institutional and regulatory environment as well as very particular historical reasons to focus on inclusion. This density of the institutional environment generates variation, which allows discovery of the finer interactions with organisational form. Furthermore, due to South Africa's long history of well-developed institutions, information on policies and regulations was well-documented and relatively easily accessible.

To collect detailed information on the different governance structures and local situation, face-to-face interviews were held on location. Such an interview-based method is particularly useful to generate insights into how and why questions (Pettigrew, 1990). Interviews were complemented with desk research for triangulation (Yin, 1994), especially on relevant laws and regulations. Interview protocols were standardized to facilitate comparisons (Miles and Huberman, 1994). Using pre-developed questions also allowed maintaining a link between existing knowledge and inductive elements of the study (Pettigrew, 1990). Questions were built on the LINK methodology, which was developed to understand inclusive trading relationships between farmers, farmer organisations, and formal markets (Lundy et al., 2012).

Interview questions could be subdivided into three main categories: (1) What is the current structure of the chain? (2) How is the current inclusivity of the chain? (3) How many inclusivity be enhanced further? Questions on the current structure focused on the chain organisation and key actors, as well as product, payment, services, and information flows. Questions on the current inclusivity identified drivers, development phases, provided support, the perceived value of this support, and success factors. Finally, questions on enhancing future inclusivity focused on willingness to invest, areas for improvement, areas with potential, and policy support. The core set of specific questions were developed by the corresponding author and adapted and complemented in the field by whoever was conducting the interviews, as the situation commanded.

Data collection for the cooperative model took place over a two-week period in October 2013 and was initiated via emails explaining the objective of our study. The first series of interviews was held by the corresponding author with the company managers responsible for developing the cooperative model. During these interviews, we learned about the existence of the third-party model. Interviews for this model, which was located in a different area, were conducted in August 2014. This second series of interviews was held by Wytske Chamberlain, a

PhD student at the University of Pretoria working on the UP-PGSARD[1] project. Interviews with managers were followed up by interviews with board members of successful and less successful cooperatives, field staff, and contractors. For the third party model, interviews were also held with managers and employees of the management company. After having interviewed the stakeholders, follow-up interviews were held with company management to clarify issues that came up. During writing up, additional contact by email and telephone was used to check details and obtain missing information.

Transaction characteristics

The basic transaction studied is the purchase of sugar cane from growers by a processing company operating a sugar mill. This company or miller is the 'driving actor' which decides - within the limits of the possible - on the arrangement used to govern the transaction. The transacted product, sugar cane, undergoes several value-adding transformations before arriving at the sugar mill. As the governance structures studied affect the scope of the transaction governed, it is instructive to comprehensively describe these transformations (Gereffi and Korzeniewicz, 1994; Porter, 1985). In this value chain, the locus lies on the arrangements millers make with growers in production, procurement, and sales of sugar cane.

Sugar cane grows from a perennial rootstock, which when properly maintained yields sugar cane for up to ten years. Soil preparation and planting only has to take place in the first year; all other activities take place per crop cycle. In KwaZulu-Natal, our study region, these crop cycles are 12 to 15 months long. Ratoon maintenance takes place during the first six months of each cycle and consists of applying fertilizer and controlling pests and weeds.[2] Harvesting is done manually, with a machete. First, the field is burned to remove excess foliage and reduce the presence of snakes and subsequently cut into stacks. Tractors are used to haul the stacks from the field to dedicated loading areas, scattered throughout the cane growing region. Here, the cane is loaded onto trucks which transport it to the sugar mill.

Governance structures economize on transaction costs, which result from transaction attributes (Williamson, 1996). In institutional economics, the most important attributes of governance structures are considered to be transaction frequency, uncertainty, and asset specificity, with the latter being particularly important (Joskow, 2005; Klein, 2005). For sugar cane purchasing, transaction frequency is low when considered as the times the miller buys from an individual grower. With production cycles of 12-15 months and small landholdings, the miller on average buys 0.8-1 times per year from each grower. However, the miller has to conduct thousands of these transactions each production cycle and thus prefers a highly standardized approach.

Transaction uncertainty can derive from volatility caused by exogenous shocks or ambiguity caused by difficulty in observing the state of the environment (Carson et al., 2006). Ménard (2012a) identifies five potential sources of uncertainty mentioned in the literature: unstable or uncertain demand, technological change, variable quality and timing of delivery, risk of opportunism, and appropriability hazards caused by inadequate institutions. In sugar cane purchasing, ambiguity and volatility in raw product quality and timing of delivery are the most important sources of uncertainty. Demand is stable and given by the capacity of the sugar mill, which is expected to remain below full capacity even when all available land is producing optimally, which was not the case at the time of research (SASA, 2014). Technological change in sugar cane sourcing is slow and unlikely to affect the nature of the transaction. Opportunism is low due to high physical asset specificity, explained in more detail below. Finally, appropriability hazards are limited to the distribution of generated value, which in the South African sugar cane sector is both highly regulated, as explained in the next section, and possible to define ex-ante.

The most important transaction uncertainty derives from variability in product quality and timing of delivery. The sugar mill operates for 8-9 months per year. During this period, it operates 24/7 and, given high fixed costs, is most profitable when running at full capacity. Keeping the mill at full capacity requires precise coordination, complicated by the fact that cane needs to be crushed within 72 hours after burning to prevent quality loss. Low quality sugar cane, defined as cane with a high fibre content, can cause the mill to jam, creating hold-up problems along the entire supply chain. Low quality cane is also caused by inadequate field management. The quality of sugar cane is costly to assess before crushing. Hence, millers prefer to control the timing of burning, harvesting, and transport to reduce uncertainty.

Asset specificity is generally considered the most important transaction characteristic driving the choice of governance structure, where the specificity refers to "the degree to which an asset can be redeployed to alternative uses and by alternative users without sacrifice of productive value" (Williamson, 1991). The share of the productive value of an asset which is sacrificed when it is deployed outside of the intended transaction is its degree of specificity. Williamson (1991) mentions six types of (relationship-specific) asset specificity: site specificity, physical and human asset specificity, brand name capital, dedicated assets, and temporal specificity. Of these, site and physical asset specificity and dedicated assets are the most important in sugar cane purchasing, although the distinction between these categories is not clear-cut. The specificity originates from the high bulk-to-value ratio of sugar cane, which results in high transport costs. The maximum distance from which cane can profitably be sourced defines the extent of the catchment area. In our case study, this maximum distance was roughly

40 kilometres.[3] Constructing a sugar mill requires substantial capital investment into specific assets. Changing activity or changing the location of the plant is prohibitively expensive. Similarly, albeit to a lesser extent, growers are locked into a relationship with a specific mill. The investment into sugar cane rootstock is effectively a sunk cost: once growers located within the catchment area decide to plant sugar cane they have to produce and deliver cane for several years to recuperate the value of the initial investment.

Institutional environment

The sugar industry is heavily regulated and relatively protected compared to other South African agricultural sectors (Kirsten et al., 2009; OECD, 2006). Strong government intervention in most agricultural markets was ended with the Marketing of Agricultural Products Act 47 of 1996; however, an exception was made for the sugar industry (Kirsten et al., 2009). This policy was justified on the grounds of protectionism and subsidies in other sugar producing countries. The domestic sugar industry is regulated through the Sugar Act of 1978 and the Sugar Industry Agreement of 2000 (NAMC, 2013). Main features of the South African sugar policy are import tariffs on sugar imports from non-SACU (Southern African Customs Union) member countries, the pivotal role of the South African Sugar Association (SASA), and the establishment of the Recoverable Value (RV) price system which determines cane prices millers have to pay to growers (Dubb, 2013; NAMC, 2013).

The domestic sugar price is protected by import tariffs from cheap imports of dumped and subsidized sugar and hence generally exceeds international prices. South African sugar companies are exempted from these tariffs and thus benefit from higher prices. Mills in SACU member countries that have surplus sugar production - which only applies to Swaziland - also benefit from these higher prices. The artificially higher price constitutes an indirect subsidy to the sector, which consequently enjoys relatively high Nominal Rates of Assistance[4] (Kirsten et al., 2009).

The South African Sugar Association (SASA) is at the heart of the industry. The board of SASA is composed of representatives of millers (South African Sugar Miller's Association) and growers (South African Cane Growers' Association). It operates independently and free of government intervention. The chairman and vice-chairman are elected every two years and rotate between representatives of millers and growers (NAMC, 2013). The mandate of SASA includes control over both domestic and international marketing of South African sugar and thus also over exports of bulk raw sugar (NAMC, 2013). It regulates the relationship between growers and millers by establishing the distribution of proceeds and setting and publishing the Recoverable Value (RV) prices for sugar cane. The

sucrose level required to establish the RV price is measured at all South African sugar mills by the Cane Testing Service, a SASA subsidiary (SASA, 2014). SASA also hosts the South African Sugar cane Research Institute (SASRI) and the Shukela Training Centre, which provides training to new sugar cane growers, established commercial and small-scale growers and farm workers (SASA, 2014). Finally, as the representative organisation of the sugar industry, it is responsible for information gathering and external communication on all industry affairs.

Sugar cane prices received by growers are controlled by the SASA using the above-mentioned RV price system, which is based on the sales of local sugar, exported sugar, and molasses. Of the proceeds remaining after deducting levies, growers receive 64% - a ratio which has not changed since 1994. This is the RV price. The actual price received by growers depends on the quality of the cane delivered to the sugar mill. Cane quality is determined by the sucrose, non-sucrose, and fibre content of each batch of cane, which is captured in the RV% (Wynne et al., 2009). This RV% is multiplied by the RV Price to determine the price to be paid to the grower (Canegrowers, 2013).[5] Hence, the RV price received by growers is fixed by SASA, yet through the RV% quality incentives are maintained.

The South African sugar industry has always had an oligopsonistic structure, with several large milling companies and thousands of large and small growers. As of 2013, there were six milling companies operating fourteen mills, producing up to 2.2 million tons of sugar per season (South African Sugar Industry, 2013). These millers sourced sugar cane from just under 24,000 registered sugar cane growers (SASA, 2014), which is only half of the number of growers registered in 2000s (Dubb, 2013). A large majority of these registered growers are smallholder farmers, of whom only half sold sugar cane in the 2014 season. The cane they supplied represented about 8% of the total crop (SASA, 2014). The remaining 1,413 registered growers are large-scale producers, of whom 323 are black emerging farmers and the rest, the large majority, white commercial famers. Together they account for 84% of the sugar cane produced. The milling estates supply the remaining 8% (SASA, 2014).

With the fall of the apartheid regime in 1994, the new ANC government under Nelson Mandela undertook an ambitious land reform policy, resting on three pillars: land restitution, redistribution, and tenure reform. Land restitution allows communities whose land was seized after 1913 to register claims to land ownership or receive financial compensation. Under land redistribution policies, estate land is acquired with financial assistance from the state and redistributed amongst previously disadvantaged groups. Tenure reform consists of legal changes to guarantee security of tenure (Ministry of Land Affairs, 1997).

Land reform has heavily affected the South African sugar industry, which obtains most sugar cane from regions with large areas of communal land, such as KwaZulu-Natal. Land reform initiatives have already resulted in the transfer of

70 627 hectares of land formerly owned by large-scale white growers and milling estates to black owners, equal to 21% of the total land area used for sugar cane (CANEGROWERS, 2013). In total, more than 50% of the freehold land under cane has been claimed within the land restitution programme. However, the pace at which claims are settled is painstakingly slow (NAMC, 2013). In recognition of this problem, in 2004 the industry established the Inkezo Land Company with the aim to 'streamline transfers of ownership by identifying sellers and buyers, streamlining processes of land reform and promoting the sustainability of the new ventures through outsourced support service providers or partners' (NAMC, 2013; South African Sugar Industry, 2013). New landholders often acquire user rights to small plots of land held under communal land tenure systems.

Sourcing from small-scale growers is nothing new; Dubb (2013) describes a history of small-scale growers starting in the 1950s. However, with land reform an increasing share of land within the catchment area of sugar mills is being managed by such small-scale growers. Most lands on which small-scale producers grow sugar cane fall under traditional community areas, in which land use and ownership are regulated through a communal land tenure system. In addition, most instances of restituted and redistributed land are also farmed on a communal basis. Such land officially remains state land, but user rights are vested in community members. One major implication is that such communal land is not freely transferable. It cannot be bought or sold to members outside of the community and user rights can only be transferred within the community by inheritance or by permission from the traditional authorities concerned. It may also be held in Communal Property Associations, which manage land acquired by a community (Republic of South Africa, 1996).

Three governance structures

The institutional context described above is important to understand the emergence of different governance structures between growers and the miller. Following Williamson (2000), the institutional environment, along with the embedded rules and customs, shapes institutional arrangements between transacting parties which align governance structures with transaction costs.

The definition of the governance structure crucially depends on how the scope of the transaction is defined. Tongaat Hulett Sugar sources sugar cane from three different sources: own plantations, medium-scale growers, and small-scale growers. Distinct arrangements are used to govern these transactions. Defining the scope as sugar cane sourcing would consider the governance structure to be a plural form, defined as 'those organisational arrangements in which, for a class of transactions dealing with the same activity and within the same institutional and competitive environment, a party uses simultaneously different modes of

Table 2.1. Control of activities and asset ownership per governance arrangement

Activities	**Market**	**Cooperative**	**Third party**
Soil preparation and planting	grower	miller	third party
Ratoon maintenance	grower	cooperative	third party
Harvesting	grower	cooperative	third party
Transport	miller	miller	miller
Assets			
Land	grower	grower	grower/coop.
Rootstock	grower	cooperative	miller
Sugar mill	miller	miller	miller

governance or relies simultaneously on substantially different types of contracts' (Ménard, 2012b). Limiting the scope to purchases from small-scale growers would still consider the governance structure to be plural, albeit at a different level, as three distinct arrangements are used to govern such purchases. In this paper, however, the three arrangements used to purchase sugar cane from small-scale growers will be analysed separately to understand which factors explain the choice for each arrangement.

Sugar cane purchases from small-scale growers are governed by three distinct arrangements, which will be referred to as Market, Cooperative and Third-party. These arrangements differ in who controls specific activities and assets. An overview is given in Table 2.1. In the Market arrangement, there are only two parties controlling the main activities and assets: the grower and the miller. Although intensive communication and coordination of activities is required, parties retain their autonomy. Growers control on-farm activities, from planting to maintenance and harvesting. The main assets in this process, land and the sugar cane rootstock, are also under the grower's control. Transport of the sugar cane to the mill is organized by the miller, albeit executed by dedicated transport companies.

The Cooperative arrangement involves an additional actor: the cooperative. Growers lease their land to the cooperative for ten years. In exchange, growers receive rental income directly and, through cooperative membership, a share of cooperative profits. Soil preparation and planting, as well as ratoon maintenance in the first year is coordinated by the miller and financed through a development grant from the provincial government. As soon as the sugar cane reaches canopy stage for the first time, around 6-7 months after planting, the ownership of the rootstock is transferred to the cooperative, which is responsible for ratoon maintenance for the remainder of the productive life of the rootstock. Harvesting is coordinated by the cooperative from the first year; transport is organized by the miller.

In the third-party arrangement, growers sign a power of attorney giving the cooperative the right to sign land-lease agreements on their behalf with the miller. In this arrangement, the miller directly owns the rootstock. Land rental payments are made to the cooperative, which distributes these rents based on the amount of land each grower contributed to the cooperative. Soil preparation, planting, ratoon maintenance and harvesting - the entire production process - is coordinated by a third party under a management contract signed with the miller. The third party is an independent management company specialised in managing sugar cane plantations, with own equipment and staff. A large number of cooperative members work as full-time employees for this company, around 6,000 at the time of writing. No additional sub-contracting is required; the third party independently manages the entire production process.

Since all three governance structures share the same institutional context, the RV price system is the same for all growers, independent of the governance structure. This does not mean that growers necessarily receive the same price, since this depends on the quality of cane delivered to the mill. In the case of the cooperative and third-party system growers depend not only on their own productivity, but also on that of other members, both for land rents and dividend income.

Choice of governance structure

Reasons to go hybrid

In the early literature on hybrid governance structures, they were identified through their differences and commonalities with markets and hierarchies (Rubin, 1978). Williamson (1991) refined this approach and showed hybrids could be considered an intermediate form between the extremes of markets and hierarchies. He explained their occurrence using the discriminating alignment hypothesis, which states governance structures align with exchange attributes to minimize transaction costs. A shortcoming of this approach is that it assumes the minimum-cost structure is always selected. In reality this is not always the case, as is evident from the existence of plural forms, where several governance structures governing identical transactions exist within the same company (Ménard, 2012b). In such a context the approach might still be valid to examine how the different governance structures affect transaction costs. That is the approach taken here.

Ten years ago, all sourcing from small-scale sugar cane growers was conducted through the market arrangement described in the previous section. However, this approach was found to have some shortcomings. In the market arrangement, intensive coordination between company extension staff and individual growers regarding the timing of the harvest and subsequent transport to the mill was required to control the uncertainty stemming from the variable quality of the

sugar cane. Because of the small size of each transaction, these coordination costs were substantial. Due to population growth and the custom of dividing land holdings equally between children - many of whom subsequently built their houses on this land - the average size of land holdings fell sharply, as did the total amount of land available for sugar cane production. With smaller landholdings, transaction sizes started to fall, and with an increasing number of producers the number of transactions increased. Coordinating this increasing number of small transactions resulted in ever-increasing transaction costs.

Variable quality was not caused only by burn-to-harvest-to-crush delays. Sugar cane requires large quantities of inputs and careful maintenance in the months before reaching canopy stage in order to produce a high quality crop. Insufficient investment caused many small-scale cane growers to produce a crop of variable, often low quality sugar cane. Under the market arrangement, company extension staff conducted field visits to detect such quality shortfalls. These visits, however, did not resolve the core of the problem: financial constraints. Limited financial means also resulted in many growers being unable to afford replanting the rootstock at the end of its ten-year productive cycle. With assets under separate ownership, the miller was unable to resolve these financial constraints without creating appropriability hazard, and growers had insufficient assets to use as collateral to access financing in the open market. Although financial constraints by themselves might be unsatisfactory explanations for hybrid models - as mentioned by Rubin (1978) - they were an important additional factor driving towards more integrated governance structures.

The high transaction costs under the market arrangement created clear incentives to move towards more integrated forms of governance; however, they do not explain why the optimal solution would not be complete vertical integration. In this case, the most important reason not to vertically integrate was institutional. As explained in Section 3, under current land laws communal land cannot be individually owned. With full integration out of the question, the required scale could only be reached by pooling resources, which implies parties should relinquish some control. Such a surrender of control will only happen when both parties have an incentive to do so (Ménard, 2012a): the value generated by bringing down transaction costs through resource pooling needs to be distributed in such a way (ex-ante rent allocation) that both parties are better off under the hybrid arrangement. The question then becomes: what specific form should such a hybrid arrangement take?

Choosing a specific hybrid form

To understand the choice for the specific form of hybrid arrangement chosen it is necessary to go beyond the drivers explaining why the transaction is suited to be

governed by a hybrid structure.[6] Ménard (2004) distinguishes three dimensions of transactions giving hybrids their 'personality'. The first is the monitoring mechanism designed to deal with uncertainty and contractual hazards stemming from mutual dependence. The second relates to the allocation and protection of rents created by the chosen arrangement. The third dimension concerns dispute resolution and contract enforcement. These dimensions will be used to describe both the cooperative and third-party arrangement, to show to what extent they explain the relative advantages and disadvantages of each arrangement. Although these dimensions do provide valuable insights into the workings of the arrangements, they are not sufficient by themselves to explain why the observed arrangements took their specific form. Hence, additional drivers, specific to the case study, will also be discussed.

The monitoring mechanism is the linchpin of both arrangements: it is the mechanism by which coordination costs - the size of which was the most important reason to switch from the market arrangement to hybrid arrangements - are controlled. Under both arrangements, more control at lower costs is attained by pooling land in production units, which increases transaction size. Coordination of harvesting takes place between representatives of these production units and company extension staff. In the cooperative arrangement, cooperative management is authorized to make decisions for all land contributed by its members. In the third-party arrangement, land is under direct control of the miller, who instructs the management company on desired time and quantity of harvesting. Grouping makes more intensive coordination feasible, effectively reducing the risk of quality losses resulting from to burn-to-harvest-to-crush delays.

The increased scale of production made possible by pooling land formerly under separate ownership generated economies of scale in production and monitoring. Cooperatives hire contractors for ratoon maintenance and harvesting. Each contractor manages a minimum of 40 hectares, which is sufficient to allow contractors to profitably invest in specialised machinery - bringing down production costs. Each cooperative hires several contractors, which are monitored regularly by cooperative management and company extension staff. Hiring several contractors simultaneously creates market pressure on prices and quality and revolving one-year contracts keep contractors on their toes. However, this system required the creation of a large number of capable local contractors, which was time and resource consuming. Hence, in the third-party arrangement, a bilateral agreement was made with a single pre-existing large contractor, which was made responsible for managing all land governed by the third-party arrangement - several thousand hectares. As a result of this agreement, the contractor had to expand its operations substantially, investing heavily in dedicated assets.

The reduced transaction costs in both hybrid arrangements created value which had to be distributed between partners in an incentive-maximising way. Each

partner had to expect sufficient returns ex-ante to participate and have limited scope for ex-post opportunistic behaviour. One of the most important components of this rent allocation mechanism is the price, which in South Africa is controlled by the government. As mentioned above, growers receive a fixed price, the RV price, which is published monthly by the South African Sugar Association and depends on the quality of the delivered cane through the RV%. However, in the hybrid arrangements the line between miller and grower becomes blurred. In both hybrid arrangements, growers receive 10% of gross proceeds as land rental. In the cooperative agreement, they additionally receive dividends as cooperative shareholders. Contractors get hired based on a tender process. In the third-party arrangement, the management company gets paid gross proceeds minus land rental and saving for re-planting. As the owner of the rootstock, the miller receives this re-planting saving. The influence of SASA and the government - albeit indirectly through the tariff system - on the reward system is substantial. By setting the price and pre-determining the distribution of proceeds between millers and growers, all parties have an incentive to produce high quality sugar cane, at least as long as the price is high enough to justify the required investment in inputs. Furthermore, this pre-determined distribution reduces the scope for post-contractual fighting over proceeds.

The mutual dependence between growers and millers inherent in the sugar cane supply base creates strong incentives for informal dispute resolution. Generally, both parties would lose from breaking the relationship and hence have an incentive to resolve disputes. However, there are still risks of opportunism. In the cooperative arrangement, this risk lies not as much in the relationship between the company and the cooperative, as in that between the cooperative management and its shareholders. Once elected, cooperative management has far more information over realized profits than individual members. To prevent opportunism, the company created a project office responsible for the financial management of the cooperatives. The project office pays land rent directly to the growers and retains earnings for next year's ratoon maintenance and re-planting into a savings account managed by SASA. Contractors are paid directly out of this savings account upon a joint request from the contractor, cooperative management, and company extension staff. Only profits are paid to a bank account controlled by cooperative management.

No such system is required for the third-party arrangement, because of the high commitment of transaction-specific assets. Although formally the miller could choose to dissolve the contract, it would be difficult to find another management company with the required scale and capabilities. At the same time, the management company has invested so heavily in transaction-specific machinery, it would incur substantial costs if the contract were terminated unexpectedly. The bilateral and balanced nature of this arrangement allows it to depend more on the

quality of the relationship than on formal structures of control.

Describing the cooperative and third-party arrangement using Ménard's (2004) three dimensions is informative, in that it generates additional insights into the functioning of the arrangements and trade-offs faced that cannot be derived only from the transaction attributes used by Williamson (1991) to analyse hybrids. However, they do not by themselves explain why the chosen arrangements took the specific form they did. For example, an important additional factor important to understand the structure of the cooperative arrangement is the presence of a government subsidy. The KwaZulu-Natal provincial government, in an attempt to create employment, agreed to finance the re-planting of several thousand hectares of sugar cane on communal lands through a grant. With the rootstock financed through the grant, it was not necessary for the company to have direct control over the land in the same way as in the third-party arrangement, which allowed a more hands-off approach. On the other hand, it required the use of local contractors, which at the inception of the arrangement did not yet exist.

Discussion

In this paper, transaction cost theory on hybrid governance modes was used to explain why the studied transaction was governed by the specific form of hybrid models observed. The discriminating alignment hypothesis of Williamson (1991), which states governance structures align with exchange attributes to minimize transaction costs, was used to understand why hybrid models had an advantage over the market. Subsequently, the three dimensions suggested by Ménard (2004) - monitoring mechanisms, rent allocation, and enforcement mechanisms - were used to delve deeper into the specific form taken by the observed hybrid models. Although these approaches did explain some characteristics of the studied arrangements, these theories were insufficiently precise to explain, let alone predict, the details of the different arrangements, their 'personalities'.

One area which deserves further attention is the limited focus on only those characteristics directly related to the transaction - frequency, uncertainty, and asset specificity. The studied governance structures could not be explained without taking interrelatedness in the production process into account. The move from the market arrangement towards hybrid arrangements was largely possible because of the value generated by economies of scale at production level. To realize these gains, investments were required which needed to be protected from appropriation risk. These factors were crucial considerations in the design of both hybrid arrangements. Williamson himself noted that focusing on transaction-specific characteristics alone constituted a partial analysis, and full understanding of asset specificity would require taking production characteristics into account (Riordan and Williamson, 1985; Williamson, 1991). In the context of our case

study, such a partial analysis would indeed have resulted in imprecise conclusions.

Another area which should be studied further is the role of financial constraints of one or several parties to the agreement in hybrid governance structures. Although financial constraints might be an unsatisfactory explanation for hybrids when it is used as the only explanation (Rubin, 1978), it was found to be of great importance in the case studied in this paper. Financial constraints not only were an important incentive to go hybrid, they also affected the specific form taken by the hybrid arrangements. The unbalanced investment required to create sufficient value for all parties to be interested in cooperation required a higher level of control for the party committing the additional resources.

The country-specific and local institutional environment were found to have an undeniable impact on the studied governance structures. The heavily regulated South-African sugar industry, with its fixed subsidized prices and representation structure limited the scope to incentivise small-scale producers through higher prices, but did offer the platform to facilitate saving for ratoon maintenance. Land reform forced millers to sell the land used for on-site sugar cane production and historical land ownership patterns caused this land to be transformed into the communal land holdings on which small-scale production currently takes place. In South Africa, employment generation is a responsibility of the local government, which therefore had the authority to become actively involved in encouraging participation of small-scale sugar cane producers in the agri-food value chain. It is unlikely the studied governance structures would have taken the specific form they have now without this particular institutional context.

An interesting facet of the case study which was not touched upon in the analysis so far is the reason for the coexistence of different arrangements to govern the same transaction. That is, the motivation to choose a plural form rather than a single optimal solution. Ménard (2012b) identified three factors as possible motivations: ambiguity surrounding the fitness of a mode of organisation to the transaction at stake, complexity of a transaction, and strategic behaviour. Of these factors, the initial ambiguity surrounding the fitness of a mode of organisation appears the most important determinant in the case studied here. Dissatisfaction with the market mode of organisation motivated development of the cooperative model. However, difficulties with finding qualified local contractors and consequent expensive training and supervision resulted in the conception of the third-party model, which avoided the need for local contractors. As these models were developed recently, time will tell whether they continue to exist in their current form, or whether governance converges to a single preferred mode.

Conclusion

The objective of this paper was to see to what extent existing theory on hybrid

governance structure could explain the specific form taken by a set of inclusive business models in sugar cane sourcing in South Africa. Although some aspects were explained, it was not possible to offer a comprehensive explanation. Transaction characteristics, the institutional environment, and Ménard's dimensions of hybrid arrangements were used to study three different arrangements governing the same transaction. Clearly, as the transaction characteristics and the institutional environment were identical for each arrangement, these factors alone were not sufficient to completely explain their 'personality'. Although Ménard's three dimensions did allow comparing and understanding these personalities, by themselves they were not specific enough to explain the finer details.

At least two areas that require further attention were identified: production characteristics of the local value chain and financial constraints of transacting parties. Production characteristics affect the scope for value creation and control. In the case of sugar cane, it determined where economies of scale could be attained and the almost complete dependency on a single buyer, limiting potential side-selling risk. Financial constraints affect the scope of possible governance structures. Since the firm is responsible for making most of the investment, in exchange it requires a certain degree of control, directly shaping the specific form of the governance structure.

The importance of the institutional environment was clearly demonstrated, highlighting in particular the importance of national and local regulations and institutions. In a way, the institutional environment is like a piece of glassware within which governance structures have to find their optimal form. It prevents some forms and enables others. At the same time, it is not inflexible. It changes its shapes over time, at least partly in direct consequence of what happens within. It is paramount to further develop theory, to make it more sensitive to the finer determinants of governance structures. The current lack of sensitivity of theory is particularly relevant in the context of dynamics. Governance structures are far from static, changing form and structure over time in response to internal and external pressures. Understanding coexistence, 'personality', and dynamics is a large and fruitful area open for exploration by future research.

Such research is warranted by the potential relevance of inclusive business models for poverty alleviation and economic development. Policy makers and international development organisations are increasingly looking for opportunities to work together with business in reaching these objectives. However, it is far from clear in what way public resources could be most effectively leveraged in this context. This case study shows how important public involvement was both in sharing risk in the development of the cooperative model, as well as in shaping the 'personality' of this model through regulations. On the other hand, the requirement to work with local contractors might risk the longevity of the model by imposing high transaction costs. Without a deeper understanding of the factors determining the

presence and personality of inclusive hybrid models, such questions will remain unanswered.

Acknowledgements

This research was based on one of ten case studies conducted for the Seas of Change Initiative and funded by the technical centre for agricultural and rural cooperation (CTA), a joint international institution of the African, Caribbean and Pacific (ACP) group of states and the European Union (EU). We would like to thank the Centre for Development Innovation (CDI) for making this research possible, with special thanks to Monika Sopov and Jan Helder; Ward Anseeuw of the University of Pretoria for logistical and intellectual support; and all the interviewees from Tongaat Hulett for their time and willingness to participate. Professor Nic Olivier from the University of Pretoria is gratefully acknowledged for his help with correctly interpreting the land reform policies and their implications. Finally, we would like to thank two anonymous referees for their helpful suggestions.

References

Anderson, K., 2009. Distortions to agricultural incentives: A global perspective 1955-2007. The World Bank, Washington D.C., USA.

Blois, K.J., 1972. Vertical quasi-integration. The Journal of Industrial Economics 20, 253-272.

Canegrowers, 2013. Newsletter of the South African Cane Growers' Association, June 2013, The Canegrower, www.sacanegrowers.co.za.

Carson, S.J., Madhok, A., Wu, T., 2006. Uncertainty, Opportunism, and Governance: The Effects of Volatility and Ambiguity on Formal and Relational Contracting. Academy of Management Journal 49, 1058-1077.

Coase, R.H., 1937. The Nature of the Firm. Economica 4, 386-405.

Dubb, A., 2013. The rise and decline of small-scale sugarcane production in South Africa. A historical perspective. Working paper 28. Institute for Poverty, Land, and Agrarian Studies (PLAAS), Belville, South Africa.

Eisenhardt, K.M., 1989. Building theories from case study research. Academy of Management Review 14, 532-550.

Eisenhardt, K.M., Graebner, M.E., 2007. Theory Building From Cases: Opportunities And Challenges. Academy of Management Journal 50, 25-32.

Gereffi, G., Korzeniewicz, M., 1994. Commodity chains and global capitalism. Praeger, Westport.

Hartley, J., 2004. Case study research, in: Cassell, C., Symon, G. (Eds.), Essential guide to qualitative methods in organisational research. Sage, London, UK, pp. 323-333.

Hodgson, G.M., 2002. The legal nature of the firm and the myth of the firm-market hybrid. International Journal of the Economics of Business 9, 37-60.

Joskow, P.L., 2005. Vertical integration, in: Ménard, C., Shirley, M. (Eds.), Handbook of New Insitutional Economics. Springer, Berlin, pp. 319-348.

Kirsten, J., Edwards, L., Vink, N., 2009. South Africa, in: Anderson, K., Masters, W.A. (Eds.), Distortions to agricultural incentives in Africa. The World Bank, Washington DC, USA.

Klein, P.G., 2005. The make or buy decision: lessons from empirical studies, Handbook of New Institutional Economics. Springer, Berlin, pp. 435-464.

Lundy, M., Becx, G., Zamierowski, N., Amrein, A., Hurtado, J.J., Mosquera, E.E., Rodriguez, F., 2012. LINK methodology: A participatory guide to business models that link smallholders to markets, CIAT Publication. Centro Internacional de Agricultura Tropical (CIAT), Cali, Colombia.

Ménard, C., 2004. The Economics of Hybrid Organisations. Journal of Institutional and Theoretical Economics JITE 160, 345-376.

Ménard, C., 2012a. Hybrid modes of organisation. Alliances, joint ventures, networks, and other 'strange animals', in: Gibbons, R., Roberts, J. (Eds.), Handbook of organisational economics. Princeton University Press, Princeton, pp. 1066-1108.

Ménard, C., 2012b. Plural Forms of Organisation: Where do we stand? Managerial and Decision Economics 34, 124-139.

Miles, M.B., Huberman, A.M., 1994. Qualitative data analysis: An expanded sourcebook. Sage, London, UK.

Ministry of Land Affairs, 1997. White paper on South African land policy. Ministry of Land Affairs, Pretoria, pp. 1-139.

NAMC, 2013. South African Sugar Industry Study. National Agricultural Marketing Council, Pretoria, South Africa.

OECD, 2006. Agricultural policy reform in South Africa, OECD Policy Brief. OECD, Paris.

Pettigrew, A.M., 1990. Longitudinal field research on change: Theory and practice. Organisation Science 1, 267-292.

Porter, M.E., 1985. Competitive advantage: Creating and sustaining superior performance. Simon and Schuster, New York.

Powell, W.W., 1996. Inter-organisational collaboration in the biotechnology industry. Journal of Institutional and Theoretical Economics 152, 197-215.

Reardon, T., Codron, J.-M., Busch, L., Bingen, R.J., Harris, C., 1999. Global Change In Agrifood Grades And Standards: Agribusiness Strategic Responses In Developing Countries. International Food and Agribusiness Management Review 2, 421-435.

Republic of South Africa, 1996. Communal Property Associations Act, in: office, P.s. (Ed.).

Riordan, M.H., Williamson, O.E., 1985. Asset specificity and economic organisation. International Journal of Industrial Organisation 3, 365-378.

Rubin, P., 1978. The theory of the firm and the structure of the franchise contract. Journal of Law and Economics 21, 223-233.

SASA, 2014. South African Sugar Association, December 2014, www.sasa.org.za.

Simon, H., 1978. Rationality as process and as product of thought. American Economic Review 68, 1-16.

Smith, A., 1776. An inquiry into the nature and causes of the wealth of nations. University of Chicago Press, Chicago, USA.

Sopov, M., Saavedra, Y., Vellema, W., Sertse, Y., Verjans, H., 2014. Is inclusive business for you? Managing and upscaling an inclusive company: Lessons from the field. Wageningen UR (University and Research Centre), Wageningen, The Netherlands.

South African Sugar Industry, 2013. Directory of the South African sugar industry, December 2014. SASA, www.sasa.org.za.

UNDP, 2008. Creating Value for All: Strategies for Doing Business with the Poor. Report of the Growing Inclusive Markets Initiative. United Nations Development Programme, New York.

Williamson, O.E., 1975. Markets and hierarchies: Analysis and antitrust implications. Free Press, New York.

Williamson, O.E., 1985. The economic institutions of capitalism. The Free Press-Macmillan, New York.

Williamson, O.E., 1991. Comparative economic organisation: The analysis of discrete structural alternatives. Administrative Science Quarterly 36, 269-296.

Williamson, O.E., 1996. The mechanisms of governance. Oxford University Press, Oxford.

Williamson, O.E., 2000. The new institutional economics: Taking stock, looking ahead. Journal of Economic Literature 38, 595-613.

Williamson, O.E., 2002. The Theory of the Firm as Governance Structure: From Choice to Contract. The Journal of Economic Perspectives 16, 171-195.

Wynne, A., Murray, T.J., Gabriel, A.B., 2009. Relative cane payment: Realigning grower incentives to optimize sugar recoveries. Proceeds of the South African

Sugar Technology Associatin 82, 50-57.

Yin, R.K., 1994. Case study research: Design and methods. Sage, Thousand Oaks, USA.

Notes

1 Project of the University of Pretoria's Post-Graduate School of Agriculture and Rural Development (UP-PGSARD), titled "Assessment of Inclusive Development Models in South African Agriculture, for the integration of smallholder farming and land reform projects in commercial value-chains".

2 After six months, sugar cane reaches 'canopy stage', at which point the foliage is so thick weeds no longer develop. There is also no more need for fertilizer applications.

3 Interview with C. Ingle, Procurement manager at THS Maidstone Mill.

4 The nominal rate of assistance (NRA) is defined as the percentage by which government policies have raised gross returns above what they would be without the government's intervention. See for example Anderson (2009).

5 The average RV price of 2012/2013 was R3,197 which gives an average cane price (at average RV% of 13%) of R389 per ton (SASA, 2014), roughly equal to 45 USD per ton (exchange rate of December 2014).

6 This limitation of attributes differentiating hierarchies from markets to explain the properties of hybrids has also been noted by other authors (Powell, 1996; Hodgson, 2002).

Chapter 3

How firm investment and appropriability hazard shape governance structure

Abstract

In this chapter, organisational economics is used to study a set of inclusive business models with diverse governance structures to identify common elements. Their particular operating context of weak public institutions and a stark financial imbalance between firm and farmer is taken as a starting point. The analysis is based on a set of ten case studies covering different agricultural products, both crops and livestock, from Burundi, Ethiopia, Kenya, Mozambique, and South Africa. These case studies were selected for their diversity, not to be representative, and were used to show how firm investment, asset specificity, and appropriability hazards affect the organisational form, in particular the safeguards contained in the governance structure. Results suggest a close relationship. Contrary to the case where parties to a transaction are able to make their own investments, in inclusive business models asset specificity reduces appropriability hazard. Alternative markets create incentives for opportunistic behaviour by enabling side-selling of inputs and outputs. The consequent appropriability hazards cause firms to adopt more extensive – and thus costly – safeguards. Appropriability hazard is reduced when farmers make credible commitments to the agreement through co-investments. Unfortunately, this option is not available for the poorest of the poor. Moreover, it seems unlikely that third-party financing may overcome this hurdle, as it does not address the main purpose of the co-investment: signalling farmer commitment.

Introduction

Inclusive business models are defined as private sector for-profit initiatives which include the poor with the stated intention to improve their well-being through mechanisms going beyond arms-length market transactions (chapter 1). Their explicit intention to deliver public goods such as poverty alleviation or food security improvement has led to substantial donor interest and a large number of reported cases (Prieto-Carrón et al., 2006; Wach, 2012). Inclusive business models come in many shapes and sizes, ranging from developing products specifically aimed at poor consumers (Prahalad, 2009) or providing employment for poor women to sell nutrient-enhanced yoghurt to the poor (Karnani, 2007). In this paper, we focus on a subset of inclusive business models, namely those models including the rural poor as suppliers or employees in the primary production stage of agricultural value chains in developing countries.

From a development perspective, this subset of inclusive business models offers tremendous promise to improve the well-being of the poor. In 2010, 60% of the rural population in developing countries was classified as poor and 34% as extremely poor (IFAD, 2010). Although academic publications on the governance structures of inclusive business models are practically non-existent, impact studies on the effect of market access, often in the form of inclusion mechanisms such as contract farming (Prowse, 2012) and plantation employment (Maertens and Swinnen, 2009), typically demonstrate positive effects on the well-being of those who are included. Therefore, if inclusive business models succeed in reaching the rural poor, the expectation is that this inclusion will positively affect their well-being.

Although improving the well-being of farmers might be of interest to the firm, this is never the only motivation for firms to work with small farmers. The model must be commercially viable: "without a business case, there is no development case" (Sopov et al., 2014). We posit that value generation in inclusive business models originates from any of three factors: supply, quality, or corporate social responsibility. Supply refers to situations where small-scale farmers are the only potential suppliers, yet under prevailing market conditions they are excluded from the market. Quality refers to situations where firms require going beyond arms-length market transaction to ensure that quality standards are met. Finally, corporate social responsibility refers to situations where value is generated because of linking with small-scale farmers, in terms of price premiums or company brand image.

If both farmer and firm stand to gain from doing business together, why are so many small-scale farmers still excluded? One of the problems is financial. Most small-scale farmers have insufficient financial resources or access to credit to make the investments required to profitably commercialise their production. One

way to overcome this issue is through having this investment made by the firm. In that case, the value creation required to make the relationship mutually beneficial relies on the extent to which the firm is willing to invest in the relationship, which in turn depends on the extent to which it can protect and appropriate the benefits from this investment.

Traditional neoclassical economics cannot explain how such a transaction should be structured. New institutional theories of agency, property rights, incomplete contracting, and Williamson's transaction cost economics offer a sharper theoretical lens by placing the analytical focus on the economic relationship itself. Agency theory focuses on how information asymmetry creates incentive incompatibility between parties, and how to design optimal reward systems to realign incentives when enforcement is costly (Fama, 1980). Such an agency problem only exist when the agent has decision rights which affect the principal's utility. Property rights theory is concerned with how the allocation of property rights affects incentives (Coase, 1960). Incomplete contracting builds on property rights theory by prescribing optimal asset ownership based on residual control rights – the right to use an asset in a situation not prescribed by the contract between parties (Hart, 1995).

Agency, property rights, and incomplete contract theory each focus on important aspects of the economic relationship, but only Williamson's transaction cost theory (1996, 2002) places the analytical focus on the governance structure itself. In particular, it explains the form of the governance structure as a result of mitigating hazards and minimizing transaction costs. By making hazard mitigation an explicit objective of the governance structure, this approach effectively encompasses the central problem addressed by agency theory: information asymmetry. This focus on the governance structure is particularly relevant because of the institutional context in which inclusive business models operate. In many developing countries, justice systems are slow and inefficient or corrupt, making public enforcement prohibitively costly (Fafchamps, 2004; North, 1990). In such contexts, appropriation of returns on investment primarily relies on the private enforcement mechanisms embedded in the chosen governance structure and their ability to mitigate appropriability hazards. Since the success of inclusive business models hinges on their ability to overcome this crucial issue, the analysis in this paper will be based on organisational economics with a particular attention to the contribution of transaction cost theory.

Our paper is organized as follows. Section 2 relies on literature on transaction costs and development studies to identify potential sources of appropriability hazards. In doing so, we show the particular role of asymmetric investments made by firms adopting inclusive business models. Section 3 presents the data, more precisely how case studies were selected and how and what information was collected. Section 4 introduces and discusses the results, showing how

appropriability hazards and firm investment are affected by the physical and institutional environment. Comparing the case studies, we show that investments are only made when hazards can be adequately controlled, with the level of required control depending on the level of investment. Section 5 concludes with some policy observations.

Theory

Literature review

The complexity of including small-scale farmers in developing countries in globalized agricultural value chains has drawn ample attention in the field of agricultural economics, both in scientific and so-called 'grey' literature. A fast-growing share of this literature is concerned with the impact of market access on the well-being of small-scale farmers, often focusing on specific interventions such as contract farming or certifications. In this chapter, the focus is not on assessing impact nor on determining which farmers are included or excluded by certain inclusive business models. The focus lies on the governance structure itself, on understanding what makes these models capable of sourcing from small-scale farmers. Hence, most attention will be given to literature concerning the governance of the link between firms and small-scale farmers.

Three approaches to investigate the firm-farmer relationship may be discerned: micro, macro, and meso. Micro approaches focus on the effect of transaction costs at the farm level. By accounting for transaction costs in farm-household models, it was shown that on-farm production decisions and hence household welfare was directly affected by transaction costs (de Janvry et al., 1991; de Janvry and Sadoulet, 2006). Empirical work distinguished between fixed transaction costs – the costs of accessing the market – and variable transaction costs – per unit costs lowering farm-gate prices (Goetz, 1992; Key et al., 2000). These costs are household-specific and when sufficiently high might exclude households from certain markets – such as input, credit, or even output markets – altogether. Governance structures going beyond arms'-length transactions might be particularly attractive to smallholders when transaction costs are high (Key and Runsten, 1999). In some situations, the access to inputs or credit included in such beyond-market governance structures are more valuable to small-scale farmers than access to higher-value output markets (Abebe et al., 2013).

Macro approaches focus on the functioning of entire markets or value chains. When public institutions are weak, particularly when justice systems are expensive and inefficient, second-best forms of market organisation occur, relying more heavily on social forms of theft control and contract enforcement (Bardhan, 1989; Fafchamps, 2004). Value chain approaches focus on the transition from

local markets to higher value markets, with higher quality requirements. Such markets may include local supermarkets and agro-industry as well as export markets. In order to reach higher value markets – for example for fresh vegetable export – these value chains need to transition in order to maintain required quality standards (Ruben et al., 2007). Efficient value chain organisation depends on transaction costs, information asymmetry, and social and cultural elements (Grover and Malhotra, 2003; Omta et al., 2001). There is a trend towards more vertically integrated forms of value chains, mainly because of increased demand for quality and traceability (Gereffi et al., 2005; Gibbon, 2001).

Meso approaches target the level between the market and the value chain as a whole and the individual farmer: the firm-farmer relationship itself. Contracts have received most attention, particularly in developed-country agriculture (Fraser, 2005; Knoeber, 1989; Paulson et al., 2010; Sykuta and Parcell, 2003). Especially in developed countries, the increasing demand for quality and traceability increases the use of contracts (Goodhue, 2011; Goodhue et al., 2010). For developing countries, research on contract structure is scarcer - the contract farming literature almost exclusively focuses on farm-level impact (see Oya (2012) and Prowse (2012) for recent reviews). However, the research that does exists (Ali and Kumar, 2015; Grosh, 1994) highlights an important difference between contract farming in developed and developing countries. In developed countries, investments are made by the party who stands to gain most from it. Firms provide specific inputs to exert control over the production process (Hueth et al., 1999). In developing countries, firms provide inputs because farmers lack the resources to purchase these inputs for themselves (Grosh, 1994; Key and Runsten, 1999).

Although the literature on including small-scale farmers in developing countries in globalized agricultural value chains is substantial, both in size and in scope, a gap exist regarding the explanation of the specific form taken by the structure governing the economic relationship between firm and farmer. Micro approaches analyse costs and benefits mainly from the farmer's point of view, macro approaches consider either the entire market or the value chain, and meso approaches are limited to contract specifications (Grosh, 1994) or the optimal degree of vertical integration (Key and Runsten, 1999). Explaining the governance structure of inclusive business models thus requires reliance on different strands in the economic literature on governance structures, whilst remaining mindful of differences originating from the specific context of developing world agriculture.

Theoretical framework

"Unlike most economists, who are interested in the structure, functioning and implications of markets, organisational economists are interested in the structure, functioning and implications of firms" (Barney and Hesterly, 2006). Organisational

economics is not a single theory, but encompasses a broad cluster of theories. Two strands of theory are particularly relevant for developing the core idea of this paper: transaction cost and positive agency theory. A third strand, management control systems (MCSs), will be used to classify the safeguards or MCSs used to manage appropriation hazards.

Transaction cost theory assumes that opportunistic agents engage in transactions in an environment of sufficient uncertainty or complexity to surpass bounded rationality (Simon, 1978). Bounded rationality implies incomplete contracting (Williamson, 1985). The risk of opportunistic behaviour reducing mutual benefits from the transaction is particularly large when transactions involve asset specificity (Joskow, 1988). Asset specificity is the difference between the value of the asset in the relationship over its value in its next best use, i.e. the relationship-specific share of the value of the asset. Higher asset specificity increases the risk of opportunistic behaviour. As soon as one party has made an investment in highly specific assets, the other party has an incentive to re-negotiate the terms of the agreement (Hart, 1995). To mitigate such risks, transactions involving high asset specificity are more likely to be governed by more integrated governance structures, which offer more opportunities for control. Other factors identified by transaction costs theory which encourage more integrated forms of governance are high uncertainty and low transaction frequency (Williamson, 1989).

Where transaction cost theory emphasises the characteristics of the transaction, in particular asset specificity, agency theory emphasises the incentive and measurement problems of the individual (Mahoney, 1992). Such agency costs arise because of information asymmetry: the actions of one party cannot be perfectly observed by the other party, yet these actions matter to realize mutual gains. In farming, information assymmetry results from the uncertainty caused by Nature (Allen and Lueck, 2003). When yields are low, this might be attributed to lack of effort by the farmer, wrong use of inputs, or a random act of Nature. Agency theory focuses on how to structure a relationship to maximize mutual gains in the presence of such behavioural uncertainty. Two branches of agency theory can be distinguished: mathematical principal-agent models and positive agency theory (Jensen, 1998). Principal-agent models assume unbounded rationality and see the firm as a nexus of contracts, neglecting distinctive features of governance structures (Williamson, 1991), while positive agency theory acknowledges that the organisational form does matter (Fama, 1980; Jensen and Meckling, 1976). Since the organisational form of inclusive business models is the focus of this chapter, positive agency theory rather than mathematical principal-agent theory was used.

The governance structure of inclusive business models

Inclusive business models go beyond arms-length market transactions. One of the most important motives for such more integrated governance structures is the failure of markets to satisfactorily handle transactions (Casson, 1984). Governance structures going beyond market transactions may take a variety of forms, ranging from hybrids to complete vertical integration (Williamson, 1991). Hybrids are "arrangements in which two or more partners pool strategic decision rights as well as some property rights, while simultaneously keeping distinct ownership over key assets, so that they require specific devices to coordinate their joint activities and arbitrate the allocation of payoffs" (Ménard, 2004). As such, they retain some of the incentives associated with the autonomy of actors operating through markets, "but provide added transaction-specific safeguards, compared with the market" (Williamson, 1996). Such safeguards are required to limit appropriability hazards originating from inter-dependence.

One of the key features which sets inclusive business models apart from most hybrid models described in the literature thus far is their over-reliance on investment by the most financially potent partner: the firm. In this case, which of the two partners should make a certain investment is no longer dictated solely by efficiency considerations. Rather, the limited financial resources at the farmer's disposal require the firm to make most or all of the investment, even when the residual control rights over these investments *de facto*, if perhaps not *de jure*, end up in the hands of the farmer.[1] In order to be willing to make such investments, the firm requires additional controls, which would not have been necessary if farmers would have been able to make the investment themselves. Importantly, as the partner making most of the investment, and usually the one initiating the relationship, the firm has the privilege of determining the terms of trade (Barrett et al., 2012), i.e. designing the governance structure.

We hypothesize that the firm will design the governance structure in such a way that the benefits of making the investment exceed the cost of controlling it. This thesis is parallel to the finding of Crocker and Reynolds (1993), who argue that economically efficient contracts balance the ex-ante cost of contract design and negotiations with the ex-post costs of opportunism and enforcement. In the weak institutional environment in which inclusive business models operate, public enforcement is costly, making firms rely more heavily safeguards: the private enforcement mechanisms incorporated in the governance structure (Fafchamps, 2004; North, 1990). The type and extent of safeguards are expected to depend on the type and size of firm investment and the appropriability hazard: the contracting hazards exposing firm investment to misappropriation.

Firm investment, asset specificity, and appropriability hazard

A distinction should be made between total firm investment and the share of this investment which is transaction-specific. When each partner to a transaction is able to make the investments over which it has residual control, asset specificity increases the risk of hold-up (Hart, 1995) or opportunistic behaviour (Joskow, 1988; Riordan and Williamson, 1985). However, when the firm has to make an investment for the farmer, such that the farmer has de facto residual control rights over the investment, the degree of asset specificity actually reduces the risk that the farmer may use this investment outside of the relationship – where its value is lower. Absent sufficiently strong public institutions to allow cost-effective public enforcement of contractual obligations, the most credible threat a firm has to punish noncompliance is discontinuation of the relationship (Grosh, 1994). When the net present value of continuing the relationship exceeds the one-off benefit of non-compliance, a contract is said to be self-enforcing (Gow and Swinnen, 2001).

The degree of asset specificity depends on the presence of alternative markets and on the redeployability of the investment in alternative uses (Williamson, 1996), which determines its opportunity cost. In the cases under review in this paper, firm investments are rarely relationship-specific. Rather, some might be product-specific. In other words, the investments generate the highest return when used for the contracted product. However, when alternative buyers for the product become active on the local market, such product-specific investments no longer are relationship-specific, since farmers now have the opportunity to sell the same product to other firms – so-called side-selling or leakage (Fafchamps, 2004).

The risk that farmers might one way or another misappropriate the investment made by the firm is what is referred to in this paper as appropriation hazard.[2] Per definition, farmers can only misappropriate the value share of the investment which is not relationship-specific, equal to the opportunity cost of the investment. If the misappropriation was easily observable, a firm might simply discontinue the relationship as soon as non-compliance took place. If that were possible, appropriability hazard would be a non-issue. In the real world, because the output uncertainty related to natural factors is inherent to agricultural production (Allen and Lueck, 2003), it is not possible for the firm to detect non-compliance from observing delivered output alone, creating measurement uncertainty (Alchian and Demsetz, 1972). Therefore, in order to encourage compliance and detect non-compliance, firms rely on safeguards contained in the governance structure.

Safeguards

In Williamsonian transaction cost theory, safeguards have received relatively little attention, which might be due to the emphasis on transaction-cost minimization

rather than hazard or uncertainty mitigation (Sykuta and Parcell, 2003) or because of a focus on adaptability rather than control (Ménard, 2012; Williamson, 1991, 2002). The categorization which does exist, by Williamson himself (1989), distinguishes three broad types: dispute resolution mechanisms, embedding transacting parties in a 'trading network', and realignment of incentives. Dispute resolution mechanisms relate to the acknowledgement that contracts are incomplete and that it is difficult and costly to have courts effectively arbitrate conflicts. In that case, there are important advantages in relying on alternative forums for dispute resolution, thus effectively complementing public by private enforcement (Williamson, 1985). Embedding transacting parties in a trading network involves re-equilibrating trading hazards, for example by switching the relationship from unilateral to bi- or multi-lateral exchange, by relying more on collective decision-making, or even by developing joint ownership, all of which facilitates the implementation of 'hostages' – with parties making credible commitments through relationship-specific investments (Williamson, 1983). Last, realigning incentives commonly involves penalties for non-compliance, which depends to a certain extent on public enforcement. When public enforcement is weak, the most important penalty for non-compliance is discontinuation of the relationship (Grosh, 1994).

Measurement uncertainty makes it impossible for the firm to perfectly observe farmer compliance from output alone. Such a 'fuzzy output' situation is hardly unique to farming. In fact, systems to manage such situations are well-described for management within organisations (Austin, 2013; Ballwieser et al., 2012; Hoenen and Kostova, 2014). In the management accounting literature, these systems are referred to as Management Control Systems (MCSs), which are increasingly studied as a package, rather than individually (Malmi and Brown, 2008), and are applied to management within and between organisations (Van der Meer-Kooistra and Scapens, 2008). Several different frameworks exist (Malmi and Brown, 2008; Merchant and Van der Stede, 2007; Ouchi, 1992; Simons, 2013), with different levels of detail and complexity. Without feigning completeness, in this paper a simpler classification will be used, which distinguishes output, behaviour, and social controls (Langfield-Smith, 2008). Output controls concern measuring and monitoring output through performance measures; behavioural controls concern rules and standardized operating procedures; social controls are informal and derived from shared norms, values, and beliefs.

Which controls are most effective in governing the firm-farmer relationship depends on characteristics of the production process influencing the degree of measurement uncertainty. Two production characteristics are deemed to be particularly relevant: output separability and task programmability (Mahoney, 1992; Sykuta and Parcell, 2003). Output separability is narrowly defined as the ability to infer an agent's (unobservable) effort from observing output. When

output separability is high, incentive alignment is effectively attained by designing a reward system based on observable output. Task programmability concerns the relationship between observable inputs and management practices. In this case, effective incentive alignment is achieved by checking and rewarding specific practices, for example through protocols formalizing which steps should be taken to complete a specific process.

Pricing is a particular type of output control, since in addition to affecting incentives it affects the distribution of value and risk between firm and farmer. Two general types of pricing exist. Prices paid to farmers can either be fixed in advance (fixed price) or determined ex-post as a residual after subtracting processors' costs from revenues (formula pricing) (Grosh, 1994). Fixed prices encourage participation by creating certainty for the farmers, but shift almost all the risk to the firm. When spot prices fall below contractually agreed prices, the firm runs the risk of large losses; when spot prices are above contract prices, farmers have an incentive to side-sell part or all of their harvest. Formula prices, by being based on the market price, do not carry such risks for the firm but shift the risk of market price fluctuations to the farmer. Discounts and premiums are additional output controls available to a firm to encourage incentive alignment.

Propositions

We summarize the review of the literature by deriving two central propositions that intend to encapsulate the core elements of our discussion above, that is: the relation between firm investment, asset specificity, appropriability hazard, and safeguards. We view inclusive business models as a particular subset of governance structures because of the imbalance in financial capacity between firm and farmer and their strong reliance on private enforcement due to the weak institutional environments in which they operate. Most investments are made by the firm but de facto controlled by the farmer, and the general-purpose nature of most of these investments, their 'redeployability', makes them vulnerable to appropriation hazards. Consequently, it is in the interest of the firm to make the transaction-specific share of firm investment as large as possible. Thus:

Proposition 1 (level): The lower the asset specificity of the firm investment, the larger the appropriation hazard, and hence higher the reliance on safeguards.

One of the most important factors determining whether an investment is relationship-specific or not is the presence of alternative markets. The possibility of side-selling, combined with asymmetric information between firm and farmer, create appropriability hazards. If appropriation hazard is large, the size of the

investment depends on the degree to which this hazard can be cost-effectively controlled.

Proposition 2 (type): The types of safeguards chosen are affected by output separability and task programmability. Ceteris paribus, the larger are output separability and task programmability, the better appropriation hazard may be controlled, and hence the larger the investment.

We now turn to our empirical sample in order to substantiate these two propositions.

Case studies

Ten case studies of inclusive business models in agricultural sourcing were selected over a variety of products and settings throughout the African continent. These case studies were selected not to be representative of the population of inclusive business models, but rather to demonstrate the variety within the population. Since the project constituted the first attempt to consider these models as a group, few a priori criteria existed. Therefore, a more explorative approach was chosen. Research and practitioner organisations in sub-Saharan Africa were contacted and asked to provide short descriptions of inclusive business models they were familiar with. It is based on these descriptions that our ten cases were selected with two main criteria: (1) that the cases were as diverse as possible regarding their governance structures; (2) that the cases included different products, including annual, perennial, (non-) perishable crops and livestock, with different end-markets, and in different parts of the continent. The countries, products, and a general description of these cases are summarized in Table 3.1.

To collect detailed information on the different governance structures and local situation, face-to-face interviews were held on location. Such an interview-based method is particularly useful to generate insights into how and why questions (Pettigrew, 1990). Interviews were complemented with desk research for triangulation (Yin, 1994). Interview protocols were standardized to facilitate comparisons (Miles and Huberman, 1994). Using pre-developed questions also allowed maintaining a link between existing knowledge and inductive elements of the study (Pettigrew, 1990). Questions were built on the LINK methodology, which was developed to understand inclusive trading relationships between farmers, farmer organisations, and formal markets (Lundy et al., 2012; Vorley et al., 2009) and complemented by knowledge accumulated within the 'international learning initiative on scaling inclusive agri-food markets'.[3] Specifically, more detailed questions were added on barriers and enablers of model success and

Table 3.1. Case studies

Country	Product	General description
Burundi	sorghum	Sorghum, a local crop, is bought for beer production, through farmer associations, model farmers, and formal contracts.
Ethiopia	passion fruit	Fruit processing plant with own plantation and the ambition to source 50% from smallholder cooperatives with shareholding.
Ethiopia	teff, wheat, maize	Seed production on own farm and from smallholders, with full-service provision. Formal contracts are signed.
Ethiopia	sesame	Hulling and export company, 'legally-binding' contracts with smallholder cooperatives which receive pre-planting credit.
Kenya	dairy	Dairy business hubs bring farmer services and collection / bulking together. Services are deducted from milk payments.
Mozambique	cassava	Mobile cassava processing unit buys at fixed price, offers transport services, and sells to local beer company.
Mozambique	fresh eggs	Joint venture of rearing and selling company, providing services and inputs for fresh egg production to local farmers.
Mozambique	broiler chickens	Chicken breeding, hatchery, feed-mill and slaughterhouse operation has local out-growers rearing the broiler chickens.
South Africa	vines and wines	Equity share scheme, where former employees became shareholders with board and management representation.
South Africa	sugarcane	Land is leased to farmer-owned cooperatives, which become production centres, using contractors for field management.

ownership and financing of key assets.

Interview questions were subdivided into three main categories: (1) What is the current structure of the chain? (2) How is the current inclusivity of the chain? (3) How may inclusivity be enhanced further? Questions on the current structure focused on the chain organisation, key actors, and product, payment, services, and information flows. Questions on the current inclusivity identified drivers, development phases, provided support, the perceived value of this support, and success factors. Finally, questions on enhancing future inclusivity focused on the willingness to invest, areas for improvement, areas with potential, and policy support. The core set of specific questions were developed by the corresponding author, and adapted and complemented in the field by whoever was conducting the interviews, as the situation commanded.

Data collection started in October 2013 in South Africa by the corresponding author. The South African case studies were used to refine the core questions and establish the overall approach. Other case studies were conducted in the period from November 2013 to January 2014 by Yeray Saavedra and Henric Verjans of the Centre for Development Innovation in the Netherlands, Yared Sertse from the African Agribusiness Academy, and several local consultants. The approach was similar for all case studies. Contacts were initiated via emails explaining

the objective of the study. Initial interviews were held with representatives from the lead firm. During these initial interviews, other stakeholders were identified and subsequently contacted and interviewed. These stakeholders included intermediaries, farmers, entrepreneurs or employees, and where relevant, government representatives and employees of non-governmental organisations. After having interviewed the stakeholders, follow-up interviews were held with company management to clarify issues that came up. Additional contact by email and telephone was used to check details and obtain missing information. Even though the cases cover a variety of crops and countries, the same core problem was mentioned by all: "how to appropriate returns from the investments required to make the inclusive business model profitable".

Results

In what follows, our two propositions were tested using the data collected on the ten case studies described in the data section. More precisely, we look at the different types of investments, safeguards, and price setting in order to illustrate the relations suggested by our theoretical discussion and the propositions we derived. We thereafter discuss these results and the inherent limits due to our small sample, which makes quantitative tests impossible at this stage. Although we stick to qualitative information, the collected data generates an insight into common drivers shared by this diverse set of cases.

Firm investment

Investment type and level in each of the case studies are shown in Table 3.2. Investment was subdivided in three categories: training, inputs, and long-term investments. Training investments usually take the form of agents regularly visiting farmers in order to share agronomic practices. Although training through farm visits is probably the most common, other forms of training exist. For example, training may take place through demonstration plots or group sessions, which can be interactive or classroom-style (Waddington et al., 2013). More training is required when the product is a newly introduced crop, but additional training is often necessary in order for farmers to meet the quality standards required by the firm, which usually exceed local standards. Input investments depend on the product. Some agricultural products, like broiler chickens, are capital-intensive, requiring significantly more transaction-specific inputs than crops like wheat or maize. Long-term investment requirements are also largely determined by the product. For example, sugarcane is a perennial product which requires investment in rootstock and can take several years before becoming productive; passion fruit requires trellises; animals require housing. Where input investments

Table 3.2. Firm investment

Country	Product	Training	Inputs	Physical assets	Total firm investment
Kenya	dairy	=	=	=	n/a
Mozambique	cassava				low
Burundi	sorghum	+	+		medium
Ethiopia	sesame	+	+		medium
Ethiopia	teff, wheat, maize	+	+		medium
South Africa	sugar cane	+	+	++	high
Ethiopia	passion fruit	++	+	+	high
Mozambique	fresh eggs	++	++	=	high
Mozambique	broiler chickens	++	++	=	high
South Africa	vines and wines	=	=	=	n/a

Note: + some investment, ++ large investment, = co-investment, n/a not applicable.

are mostly general purpose, training investments are somewhat in the middle, and investments in physical assets are mostly product-specific.

In Table 3.2, the case studies have been ordered from low to high levels of total firm investment. On the low investment extreme, we find the Kenyan dairy hubs, which are producer-owned so that no specific firm investment is required. The governance over strategic decisions and with respect to property rights is in the hands of farmers, in a cooperative-like system. Because farmers effectively make investments in themselves and the overall level of investment is low, appropriation hazards are limited. At the other end of the spectrum, we find the case of high investment by a South African vine and wine producer. In this inclusive business model, action resides within the firm. In order to overcome incentive problems and to reduce transaction costs, existing employees of the company were given the opportunity to become shareholders, with representatives on the board of directors and the management board. The firm still exists as a single integrated entity, with hierarchical ordering; however, employees have a strong incentive to become shareholders, thus benefiting from gains in their productivity. In this co-managed governance structure, even though the level of investment is high, appropriation hazards are low, since there is little room for opportunism: all investment decisions are made jointly. In these two polar cases, we observe two modes of alignment: one with low investment and autonomous parties, and the other with high specific investments and almost completely integrated parties. In both cases, effective appropriation hazards are low.

Appropriability hazard

Between the polar cases of the almost arm's-length purchasing of dairy in

Kenya and the quasi-vertical integration of vines and wine in South Africa, appropriability hazards play a more important role. For these intermediate cases, property rights and related decision rights are separated, a defining characteristic of hybrids (Ménard, 2012). These are arrangements in which the firm and the farmers maintain separate ownership and control over those productive assets that are important for the transaction. In almost all cases, when investments are made, they are primarily borne by the firm but controlled by the farmer, which is the primary source of appropriation hazards and incentive problems.

We show that appropriability hazard increases with the size of firm investment and decreases with the asset specificity of this investment. The specific levels for each of these three variables for the ten studied cases are presented in Table 3.3. Low investment implies low appropriability hazard. When the size of the investment increases, asset specificity starts to matter. In the case of Burundian sorghum, Ethiopian sesame, and teff, wheat, and maize seeds, the asset specificity stems from the premium which farmers receive because a certain characteristic of the product is valued more by the firm than by alternative buyers. Burundian farmers receive seeds for a variety of sorghum which is particularly suitable for beer. This trait is not rewarded on the local market, where sorghum is sold for direct consumption. However, the yields for this variety are also higher, especially in combination with the included package of fertilizers and training, increasing the value of the next-best option for farmers and hence causing appropriability hazard.

When no alternative buyers are present, all product-specific investment by the firm automatically becomes relationship-specific. Hence, asset specificity is high for South African sugar cane and Ethiopian passion fruit. This does not automatically imply that the full value of the investment becomes asset specific; the degree of asset specificity depends on the value of the investment in its next-best alternative

Table 3.3. Asset specificity and appropriability hazard

Country	Product	Other buyers	Total firm investment	Asset specificity	Appropr. hazard
Kenya	dairy		n/a	low	low
Mozambique	cassava		low	low	low
Burundi	sorghum		medium	medium	medium
Ethiopia	sesame		medium	medium	medium
Ethiopia	teff, wheat, maize		medium	medium	medium
South Africa	sugar cane	no	high	high	low
Ethiopia	passion fruit	no	high	high	low
Mozambique	fresh eggs		high	low	high
Mozambique	broiler chickens		high	low	high
South Africa	vines and wines	n/a	n/a	high	n/a

use. The sugar cane rootstock and training provided in South Africa are purely product-specific; only some of the variable inputs such as herbicides and fertilizer have value for other products, or may even be sold directly to third parties. In the Ethiopian passion fruit case, trellises and young plants are product-specific and, because of the absence of alternative buyers, relationship-specific. Variable inputs specific to passion fruit growing have less value outside the relationship than in the sugar cane case, since all passion fruit growers in the region supply to the same firm, whereas in South Africa several sugar cane mills are active, each with their own catchment area. In both cases, even though total firm investment is substantial, the absence of alternative buyers makes a large share of this investment relationship-specific, resulting in low appropriability hazard.

Mozambique has active and fast-growing local markets for fresh eggs and broiler chickens, making it easy for farmers to connect with alternative buyers. Moreover, firm investment is substantial, including chicken coops, feeders, chicken feed, veterinary services, day-old chicks, and intensive training. Because of the presence of alternative buyers, only a small share of these investments may be considered relationship-specific: asset specificity is low. The combination of high investment and low asset specificity creates substantial appropriability hazard. Firms should only be willing to run such risks if cost-effective safeguards are available to mitigate these hazards.

Safeguards

The various safeguards used to mitigate appropriation hazards in the inclusive business cases that we investigated are shown in Table 3.4. As already pointed out in section 2, Williamson (1989) proposed a distinction between three categories of safeguards: dispute resolution mechanisms, embedding the transaction in a 'trading network', and realignment of incentives. When this latter category is interpreted as including not only punishment for non-compliance but also the control systems used to detect non-compliance and encourage compliance, it becomes useful to sub-divide it. To do so, we based ourselves on the definitions used by Langfield-Smith (2008), who distinguished output, behavioural, and social controls.

For all ten case studies, the Williamsonian safeguards, with incentive alignment subdivided into output, behaviour, and social controls, are shown in Table 3.4. Formal dispute resolution mechanisms were not observed in any of the cases we reviewed; however, in several cases there were designated bodies responsible for information sharing which were also involved in the decision-making process. For example, in the Ethiopian passion fruit case NGO staff, farmers, and company staff regularly convened during group training sessions and weekly sales of passion fruit ensure frequent interactions between company staff and farmers. In South

Table 3.4. Safeguards

Case	Dispute resolution	Embedded transaction	Output control	Behaviour control	Social control
Kenya, dairy	DM	farmer investment	market	-	Yes
Mozambique, cassava			fixed	-	
Burundi, sorghum			fixed	low	yes
Ethiopia, sesame	IS		market + premium	medium	yes
Ethiopia, seeds	IS		market	medium	*
SA, sugar cane	IS		market	medium	yes
Ethiopia, passion fruit	IS		market or minimum	high	
Mozambique, fresh eggs	IS	farmer investment	performance	high	
Mozambique, broilers	IS	farmer investment	performance	high	
SA, vines and wines	DM	farmer investment	internal	high	

Note: IS information sharing, DM decision making, * no association, but owner of the seed business is an active member of the local community

African sugar cane, extension workers and cooperative management had weekly meetings and jointly planned harvesting schedules and maintenance activities by contractors. In Kenyan dairy and in South African vines and wine decision making was completely integrated between firm and farmers, in both cases because farmers or employees had property and hence decision-making rights.

Interpreted in the traditional sense, all inclusive business models were embedded transactions. The farmer is selling to the firm, and the firm is investing in the farmer. Given that firm investment is the standard, only the cases in which farmers are required to co-invest in the relationship are highlighted in Table 3.4. The function of these farmer investments was identical to the 'hostages' identified by Williamson (1983): to signal credible commitment. It was observed only in the integrated Kenyan dairy hubs, the South African vine and wine business, and the Mozambican fresh eggs and broilers. In the Mozambican cases, co-investment was considered a required proof of commitment, providing important safeguards against appropriation hazards. Building chicken coops according to standards set by the firm may be a significant investment for farmers, but mainly in terms of labour commitment, something which farmers could afford.

Incentive alignment was used by all inclusive business models, with most models

using a combination of output, behaviour and social controls. Although strictly speaking it is only a form of output control when it is used to incentivize effort, pricing was a relevant instrument for all cases except for the vertically integrated South African vine and wine case. Pricing was more diverse than the fixed or formula pricing suggested by Grosh (1994). Fixed pricing was used only for Mozambican cassava and Burundian sorghum. Formula pricing, although generally not in the form of a sales price minus cost formula but rather based on going market prices, was far more common. In Ethiopian sesame it was complemented by quality premiums and in Ethiopian passion fruit with minimum prices, both measures reducing the risk of side-selling. Performance pricing is only observed for the Mozambican fresh eggs and broiler cases. In these cases, substantial firm investment is directed towards medicine, feed, and litter, which are used up in the production process. These costs are recovered from the payment the company makes to the farmers at the end of the production cycle, just like in formula pricing. However, to encourage proper use of these inputs, growers receive better prices the closer their output is to the 'optimal performance curve', the maximum attainable productivity. Such a pricing scheme was feasible because of the high output separability in fresh egg and broiler production. The relatively controlled environment reduces uncertainty caused by random acts of Nature.

The most frequently used behaviour control was monitoring (column 4), although it differed in intensity between the case studies. Monitoring and training is usually conducted by the same person, making it impossible to clearly distinguish the two (Bellemare, 2010). However, regardless of the reason for the visit to the farm, it offers an opportunity to detect non-compliance. Monitoring is most intensive for labour intensive products. In Ethiopian passion fruits, company staff visits a farm at least weekly, and Mozambican fresh eggs and broiler farmers are visited daily. Although such monitoring is clearly an attempt to directly observe effort on behalf of the farmer, it is not clear whether this is necessarily related to the degree of task programmability, since protocols and checklists were only observed for those products which had high levels monitoring, not for the other products.

Finally, social controls (column 5) were observed infrequently. In four cases farmers are organized in mutual liability cooperatives. Since farmers have a financial stake in the success of the cooperative, such types of organisation offer an incentive for social control. Non-compliant members are likely to be punished at the cooperative rather than firm level. In the Ethiopian seed case, the firm owner is a very active member of the community. Although not strictly a form of social control, this might increase the perceived cost of non-compliance for farmers by adding a personal dimension.

To sum up, one of the most important factors determining whether an investment is relationship-specific or not is the presence of alternative markets. The possibility of side-selling, combined with asymmetric information between firm and farmer,

creates appropriability hazards. If appropriation hazard is large, the size of the investment depends on the degree to which this hazard can be cost-effectively controlled. The type of safeguards used to control the hazard appear to depend at least to some extent on production characteristics. In particular, higher output separability appears to lead to more intensive reliance on output controls.

Discussion

The results suggest a relatively close relationship between firm investment, asset specificity, appropriability hazard, and safeguards. Per definition, there exists a direct relationship between the asset specificity of firm investment and the appropriation hazard. High asset specificity, ceteris paribus, reduces appropriability hazard. In line with proposition 1, reliance on safeguards was highest for the two inclusive business models with low asset specificity and high appropriation hazard – fresh eggs and broilers in Mozambique. The reverse however, was not true. When asset specificity was high and appropriability hazard low, as in the case of Ethiopian passion fruit and South African sugar cane, there was still a substantial reliance on safeguards. Independent of the asset specificity and appropriability hazard, safeguards appear to increase with the size of firm investment.

Comparing the types of safeguards with the production characteristics of output separability and task programmability provides some support for proposition 2. In particular, the relatively high output separability in the fresh eggs and broiler cases appears to be a factor in the reliance of the firm on performance pricing to align incentives. Such a clear relationship does not appear to exist between task programmability and behaviour controls, in particular monitoring. However, the intensity of monitoring does increase with the value of production. Products like passion fruit and animals might be more sensitive to quality or food safety concerns and hence justify more intense control. This corresponds with findings in a case study on several contract farming arrangements of tomato in the United States of America[4] (Hueth et al., 1999; Wolf et al., 2001).

Based on the results in the tables, an additional observation may be made regarding the relationship between the presence of alternative markets and firm investment. When alternative markets are present, firm investment is generally low. The presence of alternative markets increases appropriability hazards by creating opportunities for side-selling, and hence cause firms to avoid making substantial investments. Without alternative markets – when firms enjoy monopsony power, appropriability hazards fall substantially and investments increase accordingly. Firms are only willing to invest substantially in a lively market when these investments offer sufficient returns to warrant the intensive monitoring required to control appropriability hazard.

In addition to monitoring, firms operating under these conditions invest a lot of

time in farmer selection, use output controls, and require farmers to co-invest in production facilities, thus combining the different tools of governance that we have identified. Pushing the reasoning further, we would argue that high returns are a determinant factor because the safeguards needed to reduce appropriation hazards are very costly, particularly due to the context of weak institutions that cannot provide adequate enforcement mechanisms.

We are aware, of course, that the correlations we have established are based on a very limited sample, notwithstanding the efforts we made to select a diverse set of inclusive business cases based on the products they cover and the institutional environment in which they operate. In order to switch from a qualitative analysis to quantitative tests, a much more extensive set of cases is needed. This is an important objective, but collecting extensive information in that perspective requires major research investment, particularly given contexts that provide little institutional support and poor data. Nevertheless, we are convinced that our cases already deliver important lessons regarding the difficulties that inclusive business models raise.

Conclusion

Inclusive business models are rapidly becoming more numerous, due to their dual focus on business profitability and poverty reduction, which makes them attractive to business and development organisations alike. Organisational economics allows gaining deeper insight into their function and form. In this study, two theory-derived propositions were tested with evidence from ten case studies of inclusive business models in sourcing different agricultural products in environments with weak public enforcement. Results indicate a positive relationship between firm investment and safeguards: more investment implies more safeguards, apparently independent of the degree of asset specificity. The specific types of safeguards chosen appear to correspond with production characteristics. Output separability and task programmability facilitate control. When alternative markets are present, firm investment is limited for products with which farmers are already familiar. For high-value products with a local market, safeguards are high and farmer co-investment is required.

In a report to the FAO, DaSilva (2005) concluded that inclusive contract farming is equitable only when alternative markets are available. Paradoxically, our results suggest that the firm investment required for the inclusive business model to be developed in the first place is most likely when alternative markets are not available. This finding can be explained by the combination of asymmetric financial capacity between firm and farmer, appropriability hazards, and the context of weak public enforcement. In this context, investment relies heavily on the firm, and accompanying appropriability hazards have to be mitigated through

the governance structure. Broadly speaking, firms have two options for mitigation: incentive-aligning safeguards or demanding co-investment by the farmer. Incentive-aligning safeguards are costly, and only applied when the product is sufficiently high value, or alternative markets are absent. Co-investment is a useful complementary safeguard, but requires the farmer to have the resources to invest in the first place. The poorest farmers are unlikely to be able to meet this requirement, and thus risk being excluded. Moreover, this problem is hard to overcome by third-party financing, as it is not the investment per se but the commitment it signals that matters.

The ideal would be to develop more efficient public enforcement mechanisms, such as well-defined property rights, contract law, and efficient courts. However, building adequate institutions is a long-term and uncertain process. In the short term, the success of inclusive business models depends on the ability of parties to develop cost-effective governance structures with mutually acceptable safeguard mechanisms. There might be an important role to play by the public sector, but this is unlikely to be the direct financing of (part of) the investment, as the real problem is not lack of resources, but lack of ability to credibly signal commitment. Rather, the most important role for the public sector will most likely lie in the effective dissemination of information on best practices.

References

Abebe, G.K., Bijman, J., Kemp, R., Omta, O., Tsegaye, A., 2013. Contract farming configuration: Smallholders' preferences for contract design attributes. Food Policy 40, 14-24.

Alchian, A.A., Demsetz, H., 1972. Production, information costs, and economic organisation. The American economic review, 777-795.

Ali, J., Kumar, S., 2015. Understanding the contract structure for mango and empirical analysis of its determinants. British Food Journal 117, 2161-2181.

Allen, D.W., Lueck, D., 2003. The theory of the farm.

Austin, R.D., 2013. Measuring and managing performance in organisations. Addison-Wesley.

Ballwieser, W., Bamberg, G., Beckmann, M., Bester, H., Blickle, M., Ewert, R., Feichtinger, G., Firchau, V., Fricke, F., Funke, H., 2012. Agency theory, information, and incentives. Springer Science & Business Media.

Bardhan, P., 1989. The economic theory of agrarian institutions. Oxford University Press, New York, US.

Barney, J.B., Hesterly, W., 2006. Organisational economics: Understanding the relationship between organisations and economic analysis, in: Clegg, S.R., Hardy, C., Lawrence, T.B., Nord, W.R. (Eds.), The SAGE handbook of organisation

studies. SAGE Publications Ltd., London, UK, pp. 111-148.

Barrett, C.B., Bachke, M.E., Bellemare, M.F., Michelson, H.C., Narayanan, S., Walker, T.F., 2012. Smallholder Participation in Contract Farming: Comparative Evidence from Five Countries. World Development 40, 715-730.

Bellemare, M.F., 2010. Agricultural extension and imperfect supervision in contract farming: evidence from Madagascar. Agricultural Economics 41, 507-517.

Casson, M., 1984. The theory of vertical integration: a survey and synthesis. Journal of Economic Studies 11, 3-43.

Coase, R.H., 1960. Problem of social cost, the. Journal of Law and Economics 3, 1.

Crocker, K.J., Reynolds, K.J., 1993. The Efficiency of Incomplete Contracts: An Empirical Analysis of Air Force Engine Procurement. The RAND Journal of Economics 24, 126-146.

DaSilva, C.A., 2005. The growing role of contract farming in agri-food systems development: Drivers, theory and practice. FAO, Rome.

de Janvry, A., Fafchamps, M., Sadoulet, E., 1991. Peasant Household Behaviour with Missing Markets: Some Paradoxes Explained. The Economic Journal 101, 1400-1417.

de Janvry, A., Sadoulet, E., 2006. Progress in the Modeling of Rural Households' Behavior under Market Failures, in: de Janvry, A., Kanbur, R. (Eds.), Poverty, Inequality and Development. Springer US, pp. 155-181.

Fafchamps, M., 2004. Market institutions in Sub-Saharan Africa : theory and evidence / Marcel Fafchamps. MIT Press, Cambridge, Mass. :.

Fama, E.F., 1980. Agency Problems and the Theory of the Firm. Journal of Political Economy 88, 288-307.

Fraser, I., 2005. Microeconometric analysis of wine grape supply contracts in Australia. Australian Journal of Agricultural and Resource Economics 49, 23-46.

Gereffi, G., Humphrey, J., Sturgeon, T., 2005. The governance of global value chains. Review of International Political Economy 12, 78-104.

Gibbon, P., 2001. Upgading primary production: A global commodity chain approach. World Development 29, 345-363.

Goetz, S.J., 1992. A Selectivity Model of Household Food Marketing Behavior in Sub-Saharan Africa. American Journal of Agricultural Economics 74, 444-452.

Goodhue, R.E., 2011. Food Quality: The Design of Incentive Contracts, in: Rausser, G.C., Smith, V.K., Zilberman, D. (Eds.), Annual Review of Resource Economics, Vol 3. Annual Reviews, Palo Alto, pp. 119-140.

Goodhue, R.E., Mohapatra, S., Rausser, G.C., 2010. Interactions Between Incentive Instruments: Contracts and Quality in Processing Tomatoes. American Journal of

Agricultural Economics 92, 1283-1293.

Gow, H.R., Swinnen, J.F.M., 2001. Private Enforcement Capital and Contract Enforcement in Transition Economies. American Journal of Agricultural Economics 83, 686-690.

Grosh, B., 1994. Contract Farming in Africa: an Application of the New Institutional Economics. Journal of African Economies 3, 231-261.

Grover, V., Malhotra, M.K., 2003. Transaction cost framework in operations and supply chain management research: theory and measurement. Journal of Operations management 21, 457-473.

Hart, O., 1995. Firms, contracts and financial structure. Oxford University Press, Oxford, UK.

Hoenen, A.K., Kostova, T., 2014. Utilizing the broader agency perspective for studying headquarters-subsidiary relations in multinational companies. Journal of International Business Studies 46, 104-113.

Hueth, B., Ligon, E., Wolf, S., Wu, S., 1999. Incentive instruments in fruit and vegetable contracts: input control, monitoring, measuring, and price risk. Review of agricultural economics 21, 374-389.

IFAD, 2010. Rural poverty report 2011. International Fund for Agricultural Development, Rome, Italy.

Jensen, M.C., 1998. Organisation theory and methodology, in: Jensen, M.C. (Ed.), Foundations of organisational strategy. Harvard University Press, Cambridge, US, pp. 319-339.

Jensen, M.C., Meckling, W.H., 1976. Theory of the firm: Managerial behavior, agency costs and ownership structure. Journal of financial economics 3, 305-360.

Joskow, P.L., 1988. Asset specificity and the structure of vertical relationships: empirical evidence. Journal of Law, Economics, and Organisation 4, 95-117.

Karnani, A., 2007. The mirage of marketing to the bottom of the pyramid: How the private sector can help alleviate poverty. California Management Review 49, 90-111.

Key, N., Runsten, D., 1999. Contract Farming, Smallholders, and Rural Development in Latin America: The Organisation of Agroprocessing Firms and the Scale of Outgrower Production. World Development 27, 381-401.

Key, N., Sadoulet, E., Janvry, A.D., 2000. Transactions Costs and Agricultural Household Supply Response. American Journal of Agricultural Economics 82, 245-259.

Knoeber, C.R., 1989. A real game of chicken: contracts, tournaments, and the production of broilers. Journal of Law, Economics, and Organisation 5, 271-292.

Langfield-Smith, K., 2008. The relations between transactional characteristics,

trust and risk in the start-up phase of a collaborative alliance. Management Accounting Research 19, 344-364.

Lundy, M., Becx, G., Zamierowski, N., Amrein, A., Hurtado, J.J., Mosquera, E.E., Rodriguez, F., 2012. LINK methodology: A participatory guide to business models that link smallholders to markets, CIAT Publication. Centro Internacional de Agricultura Tropical (CIAT), Cali, Colombia.

Maertens, M., Swinnen, J.F.M., 2009. Trade, Standards, and Poverty: Evidence from Senegal. World Development 37, 161-178.

Mahoney, J.T., 1992. The choice of organisational form: vertical financial ownership versus other methods of vertical integration. Strategic Management Journal 13, 559-584.

Malmi, T., Brown, D.A., 2008. Management control systems as a package—Opportunities, challenges and research directions. Management accounting research 19, 287-300.

Ménard, C., 2004. The Economics of Hybrid Organisations. Journal of Institutional and Theoretical Economics JITE 160, 345-376.

Ménard, C., 2012. Hybrid modes of organisation. Alliances, joint ventures, networks, and other 'strange animals', in: Gibbons, R., Roberts, J. (Eds.), Handbook of organisational economics. Princeton University Press, Princeton, pp. 1066-1108.

Merchant, K.A., Van der Stede, W.A., 2007. Management control systems: performance measurement, evaluation and incentives. Pearson Education.

Miles, M.B., Huberman, A.M., 1994. Qualitative data analysis: An expanded sourcebook. Sage, London, UK.

Mujawamariya, G., D'Haese, M., 2012. In search for incentives to gum arabic collection and marketing in Senegal: Interlocking gum trade with pre-finances from traders. Forest Policy and Economics 25, 72-82.

North, D.C., 1990. Institutions, institutional change, and economic performance. Cambridge University Press, Cambridge, UK.

Omta, S., Trienekens, J., Beers, G., 2001. Chain and network science: A research framework. Journal on Chain and Network Science 1, 1-6.

Ouchi, W.G., 1992. A conceptual framework for the design of organisational control mechanisms. Springer.

Oya, C., 2012. Contract farming in Sub-Saharan Africa: A survey of approaches, debates and issues. Journal of Agrarian Change 12, 1-33.

Paulson, N.D., Katchova, A.L., Lence, S.H., 2010. An empirical analysis of the determinants of marketing contract structures for corn and soybeans. Journal of Agricultural & Food Industrial Organisation 8.

Pettigrew, A.M., 1990. Longitudinal field research on change: Theory and practice. Organisation Science 1, 267-292.

Prahalad, C.K., 2009. Fortune at the Bottom of the Pyramid: Eradicating Poverty through Profits, 2nd edition ed. Wharton School Publishing, Upper Saddle River, NJ.

Prieto-Carrón, M., Lund-Thomsen, P., Chan, A., Muro, A., Bhushan, C., 2006. Critical perspectives on CSR and development: What we know, what we don't know, and what we need to know. International Affairs 82, 977-987.

Prowse, M., 2012. Contract farming in developing countries - A review, in: Peccoud, R. (Ed.), A Savoir. Imprimerie de Montligeon, France.

Riordan, M.H., Williamson, O.E., 1985. Asset specificity and economic organisation. International Journal of Industrial Organisation 3, 365-378.

Ruben, R., Van Boekel, M., Van Tilburg, A., Trienekens, J., 2007. Tropical Food Chains. Wageningen Academic Publishers, Wageningen, NL.

Simon, H., 1978. Rationality as process and as product of thought. American Economic Review 68, 1-16.

Simons, R., 2013. Levers of control: How managers use innovative control systems to drive strategic renewal. Harvard Business Press.

Sopov, M., Saavedra, Y., Vellema, W., Sertse, Y., Verjans, H., 2014. Is inclusive business for you? Managing and upscaling an inclusive company: Lessons from the field. Wageningen UR (University and Research Centre), Wageningen, The Netherlands.

Sykuta, M., Parcell, J., 2003. Contract structure and design in identity-preserved soybean production. Review of Agricultural Economics 25, 332-350.

Van der Meer-Kooistra, J., Scapens, R.W., 2008. The governance of lateral relations between and within organisations. Management Accounting Research 19, 365-384.

Vorley, B., Ferris, S., Seville, D., Lundy, M., 2009. Linking worlds: New business models for sustainable trading relations between smallholders and formalized markets, 'New Business Models for Sustainable Trading Relations' project, (online).

Wach, E., 2012. Measuring the inclusivity of inclusive business, IDS Practice Paper 9. Institute of Development Studies (IDS).

Waddington, H., Snilstveit, B., Hombrados, J.G., Vojtkova, M., Phillips, D., Davies, P., White, H., 2013. Farmer field schools for improving farming practices and farmer outcomes in low- and middle-income countries: A systematic review, in: 3ie (Ed.). International Initiative for Impact Evaluation (3ie), London, UK.

Williamson, O.E., 1983. Credible commitments: Using hostages to support exchange. The American Economic Review 73, 519-540.

Williamson, O.E., 1985. The economic institutions of capitalism. The Free Press-Macmillan, New York.

Williamson, O.E., 1989. Transaction cost economics, in: Schmalensee, R., Willig, R. (Eds.), Handbook of Industrial Organisation. Elsevier Science B.V., Amsterdam, the Netherlands, pp. 136-182.

Williamson, O.E., 1991. Comparative economic organisation: The analysis of discrete structural alternatives. Administrative Science Quarterly 36, 269-296.

Williamson, O.E., 1996. The mechanisms of governance. Oxford University Press, Oxford.

Williamson, O.E., 2002. The Theory of the Firm as Governance Structure: From Choice to Contract. The Journal of Economic Perspectives 16, 171-195.

Wolf, S., Hueth, B., Ligon, E., 2001. Policing mechanisms in agricultural contracts. Rural sociology 66, 359-381.

Yin, R.K., 1994. Case study research: Design and methods. Sage, Thousand Oaks, USA.

Notes

1 Another way of putting this is saying that inefficiencies in the credit market – farmers are unable to obtain credit even when the investment is profitable – spill over into the product market (Bardhan, 1981). The difference with interlocked contracts, as described by Mujawamariya and D'Haese (2012) among others, is the intensive reliance on formal safeguards contained in the governance structure.

2 This definition of appropriability hazard is different from the one employed in the organisational economics literature, where it is taken to refer to "contracting hazards exposing intellectual property to expropriation" (Mayer and Salomon, 2006; Oxley, 1997).

3 This initiative is a long-term applied research, innovation and exchange program helping businesses, governments, and NGOs to share best practices (www.seasofchange.net).

4 Written out completely to avoid confusion with America, the continent, or the United States of Mexico.

Chapter 4

Verifying validity of the Household Dietary Diversity Score

based on

Vellema, W., Desiere, S., D'Haese, M. *Verifying validity of the Household Dietary Diversity Score: An application of Rasch modelling.*

Accepted for publication in the March 2016 issue of the Food and Nutrition Bulletin

Abstract

In the wake of the Millennium Development Goals, indicators have become an increasingly important tool to evaluate development projects. One of the most popular indicators to measure food security is the Household Dietary Diversity Score (HDDS). Previous research has established a link between dietary diversity and food security, however, the HDDS in the form developed by the FANTA project – 12 food groups, 24-hour recall – has never been evaluated. In this chapter, the construct validity of the HDDS is verified. A Rasch model was used to test the extent to which the HDDS meets the criteria required for interval scale measurement, using data from over 1,000 farm households in Colombia and Ecuador. Different dietary patterns between Colombia and Ecuador and two cultural groups within Ecuador required data to be split into three subgroups. For each subgroup the food groups meeting the criteria and their difficulty ranking were different. Refined indices containing only those food groups meeting the criteria contained seven items in Colombia, ten for Kichwa households in Ecuador, and nine for migrant households. These results imply that the indicator in its current form does not meet all criteria. Even when the indicator is split into refined indicators for culturally homogenous subgroups within a small region, the components of the indicator do not form a reliable way of measuring the latent variable: household food security.

Introduction

While the definition of food security formed at the 1996 world food summit (FAO, 1996) is widely adopted, disagreement remains on the indicators that assess, quantify and qualify food security and on how to operationalize these indicators at national, household or individual level (Jones et al., 2013; Leroy et al., 2015; Pinstrup-Andersen, 2009; Webb et al., 2006). Food security is measured in different ways. For example, anthropometric measures are used to monitor growth of children under five (Pinstrup-Andersen, 2009). Recalls of food consumed in the past 24hours or over a longer reference period are recorded to measure intake of macro- and micronutrients (Kennedy et al., 2010). And data on food expenditure is used to define food poverty lines (Rose and Charlton, 2002). While experience-based responses such as the Household Food Insecurity Access Score (HFIAS) elicit perceived consequences of not having enough food (Jones et al., 2013). Research institutions and development organisations alike apply such indicators to identify food insecure households or analyse effects of interventions on food security (Jones et al., 2013).

The Household Dietary Diversity Score (HDDS) is a frequently used indicator of food security. It was developed as a quick-to-implement and easy-to-use survey-based indicator to measure the impact on household food access of programs with improvements in food security as their core objective (Swindale and Bilinsky, 2006). The second version of the accompanying guide mentions that "An increase in the average number of different food groups consumed provides a quantifiable measure of improved household food access. In general, any increase in household dietary diversity reflects an improvement in the household's diet" ((Swindale and Bilinsky, 2006), p6). This suggests that the HDDS might be used as a household-level indicator of food security – indeed, it is frequently used as such (Leroy et al., 2015). However, the validity of the HDDS has never been verified, making it impossible to substantiate claims that it is a useful indicator of food security. The objective of this paper is to fill this glaring gap.

Rasch models were used to verify the construct validity of the HDDS. These models were specifically developed to test whether an additive scale consisting of several items measuring a single underlying construct meets the criteria required for interval scale measurement (Rasch, 1960). This approach differs from most other statistical techniques in that it starts from a mathematical model which meets the required criteria and tests the extent to which the data fits the model. When the data does not fit the model, it is not the model but the data which is considered wrong. By assessing the deviations of the HDDS from the criteria, specific shortcomings of the indicator can be highlighted. In effect, Rasch analysis provides the lens through which we look at the internal functioning of the indicator. Applying this methodology to analyse the construct validity of the

HDDS is the main contribution of this paper to the literature.

Household Dietary Diversity Scores

Dietary diversity refers to the variety of foods consumed by individuals or households (Jones et al., 2013; Ruel, 2003). An indicator of dietary diversity is a particularly interesting way to measure food security, because it is simple to implement, can be administered at household and individual level, and is a useful outcome in itself (Hoddinott and Yohannes, 2002). There is a shortage of validity studies of survey-based dietary diversity indicators, especially regarding the way questions are posed and how these are handled and interpreted (Leroy et al., 2015; Ruel, 2003). Particularly pressing issues are the responsiveness of food security indicators to improved food security, their discriminatory power in distinguishing food secure from food insecure households, and their validity across different cultural settings.

When measured at an individual level, dietary diversity scores are generally found to be a good proxy for micronutrient adequacy (Arimond and Ruel, 2004; Arimond et al., 2010; Hatloy et al., 1999; Kennedy et al., 2007; Moursi et al., 2008; Steyn et al., 2006). Dietary diversity might not only be linked to dietary quality, but also imply dietary quantity. According to Bennett's Law, as people become wealthier they switch from starch-dominated diets to more varied diets including vegetables, fruit, dairy products, and meat (Bennett, 1941). Although calorie intake might not increase above a certain level of wealth, Jensen and Miller (2010) suggest people quickly shift to improving the taste of their food bundle when their incomes increase. Their findings are in line with classic theories of demand (Maslow, 1943). In other words, households with sufficiently diverse diets can be assumed to at least consume enough food not to be hungry. Studies confirm a positive relationship between household dietary diversity and household food security (Faber et al., 2009; Heady and Ecker, 2013; Hoddinott and Yohannes, 2002; Kennedy et al., 2010). However, these studies were based on indicators differing in regard to their inclusion of individual foods versus food groups, number of food groups, weights, and recall period, making it hard to establish a definitive link. In fact, some authors even question what it is that is being measured by these indicators (Heady and Ecker, 2013; Ruel, 2003). In economic terms, if HDDS increases with increased food security, one should expect all food groups to increase with food consumption levels. With increasing food consumption and income, the HDDS should increase. Hence, HDDS will not be able to account for Giffen goods which have positive price elasticities.

In particular, only two research papers are named on which the conclusion that "an increase in dietary diversity is associated with socio-economic status and household food security" is based (FAO, 2012). In one of these papers, Hatloy et al.

(1999), in a case study in a southern county of Mali, indeed find such an association for socio-economic status. For nutritional status, the association was only found in urban areas. Furthermore, their index for dietary diversity is based on ten food groups, not the suggested twelve. Perhaps the most extensive work on this topic is by Hoddinott and Yohannes (2002), who study the relationship between dietary diversity and a range of food security measures using datasets covering both rural and urban households from 10 poor- or middle-income countries. The authors find a robust positive relationship – independent of whether individual foods or food groups are used to measure dietary diversity – which holds over urban and rural areas, seasons, and recall period. However, in neither of these studies is the HDDS indicator used in the form promoted in the guidelines.

Dietary diversity is measured by counting the number of foods or food groups consumed over a certain reference period. These groups can be simply counted or a weight can be attached to them based on their nutritional value. Some indicators also take into account the frequency at which the foods were consumed, or specify a minimum portion size required for a food to be counted in the index (see Ruel (2003) and Leroy et al. (2015) for a review of different indicators). Of the food-group indicators, the HDDS analysed in this paper is probably the most widely used by development organisations. It was developed by the Food and Nutrition Technical Assistance (FANTA) and actively promoted by USAID. Moreover, this index is the basis for the recent FAO "Guidelines on measuring household and individual dietary diversity" (FAO, 2012).

The HDDS was developed to measure household food access and designed to be an easy-to-use and quick-to-implement index, making it ideal for impact evaluations of development programs (Swindale and Bilinsky, 2006). It measures dietary diversity by counting the number of food groups that were consumed by the household over the last 24 hours. The indicator consists of twelve food groups: cereals; roots and tubers; vegetables; fruits; meat, poultry, and offal; eggs; fish and seafood; pulses, legumes, and nuts; dairy products; oils and fats; sugar and honey; and miscellaneous, such as condiments. These twelve food groups are based on the groups used to construct the FAO's food balance sheets (Swindale and Ohri-Vachaspati, 2005). The value of the HDDS equals the number of food groups consumed in the last 24 hours. A higher score should reflect higher dietary diversity and hence better household food access (Swindale and Bilinsky, 2006).

This study is the first to evaluate the validity of the HDDS in the form promoted in the FANTA guidelines. We limit ourselves in scope to evaluating the construct validity of the indicator, i.e. whether the different food groups contribute to a single underlying construct in such a way that the overall score on the indicator can be interpreted as an interval scale measure at household level. We do not analyse whether the scale indeed measures household food access but follow the indicator guidelines in assuming that it does. In other words, we do not study

what is measured by the HDDS, but verify how it measures.

Data

The construct validity of the HDDS was tested using data obtained from the baseline from a cross-border agricultural development project in Colombia and Ecuador. Basing our study on such 'real' data, rather than data collected primarily for research purposes, makes the results of our study more realistic. The HDDS was tested the way it is used in reality, not the way it should be used according to the guidelines.[1] These countries are culturally close and economically similar. Both countries are considered upper-middle income countries according to the World Bank classification, yet have high inequality and poverty rates. Data was collected in the Ecuadorian Amazon basin and the southern mountain range in Colombia, which are among the poorest parts of the countries. In the Amazon basin 59.7% of the population lives below the national poverty line (INEC, 2006); in Colombia's southern Andes, 50.6% of the population lives below the national poverty line (DANE, 2011).

Interviews were conducted in April and May 2012 by trained local enumerators using a structured questionnaire developed by the Centro International de Agricultura Tropical (CIAT), a CGIAR centre based in Cali, Colombia. Based on this questionnaire, a detailed protocol was developed which was used to train the enumerators during a two-week training course including field trials. Interviews were conducted with pen and paper. Raw data was cleaned and operationalized by the corresponding author, in cooperation with Martha Del Rio and Alexander Buritica Casanova from CIAT.

All interviewed households were small-scale farmers, depending on agricultural production for most of their income. Respondents were either the head of the household or the person most closely related to the head of the household, like a spouse. Since the data was collected for the baseline of a development project, the households needed to be representative for both project beneficiaries and control group households. In Ecuador, the number of surveys required to ensure representativeness of the control group was calculated based on the number of inhabitants per canton, which were obtained from the national institutes of statistics (INEC). Sampling of project beneficiaries was also done by stratification at cantonal level. In Colombia, the required size of the control group was calculated based on member lists of the national federation of coffee producers (FEDECAFE). Stratification was done at municipal level, which corresponds to the cantonal level in Ecuador, i.e. the administrative level below province (which are called departments in Colombia). Project beneficiaries were also selected for interview based on stratification at municipal level. In total, 510 households were interviewed in Colombia, and 514 in Ecuador. After removing observations for

non-response, the full dataset contained 509 Colombian and 506 Ecuadorian households.

Interviews were conducted according to a detailed standardized protocol; enumerators received two weeks of training including field trials before starting data collection. Data was collected on family composition, including ethnicity of household members[2], and income. A complete list of the definitions of the variables used is provided in Appendix B. Agricultural production destined for own consumption was valued at farm-gate prices. The used HDDS surveys were made more specific for each country by adding commonly consumed foods to the specification of the food groups (appendix C). For example, food group 1, cereals, was specified for the Ecuador survey as 'In the last 24 hours, did you consume any kind of cereal like rice, maize, or wheat, or any product made from cereals, such as bread, cookies, humitas, etc?'[3]. For Colombia, this question was specified as 'In the last 24 hours, did you consume any kind of cereal like rice, maize, or wheat, or any product made from cereals, such as bread, arepas, envueltos de choclo, noodles, puff pastries, toast, cakes, or any other food made from millet, sorghum, maize, rice, wheat, barley, oats, etc.?'[4]

Descriptive statistics are shown in Table 4.1. For the analysis, the data from Ecuador had to be split into two cultural groups, Kichwa and migrant households, as will be explained in the results section. For legibility, these groups are represented separately in the table. The first row in the table shows the months of adequate household food provisioning (MAHFP), another FANTA food security indicator. It asks for each month (recall) whether the household had insufficient food. Results are comparable between the three groups: households did not have sufficient food for two months of the year. Family sizes differed between the three groups, mostly caused by differences in the number of children, reflected by higher dependency ratios for the Kichwa and Immigrant households in Ecuador

Table 4.1. Descriptive statistics

	Colombia	**Ecuador**	
		Kichwa	**Migrants**
MAHFP	2.00	1.99	1.96
Family size	4.12	6.26	4.84
Dependency	0.40	1.27	0.70
Income (USD)*	5939	1331	2196
	(7612)	(1699)	(2040)
HDDS	8.06	5.26	6.80
	(1.52)	(1.88)	(1.80)
n	509	209	297

Note: mean and (standard deviation) of selected variables. *values for Colombia converted from Colombian Pesos using exchange rate of 31 May 2012

compared to households in Colombia. Incomes also differed starkly between the groups. Household incomes in Colombia averaged almost 6000 USD, compared to 2200 USD for migrants and a mere 1300 USD for Kichwa households. Dietary diversity, as measured by the HDDS, also differed between the groups. Colombian households scored highest, consuming on average 8 out of the 12 groups; Kichwa households scored the lowest, consuming less than half of the food groups; migrant households were in between these two groups, consuming almost 7 groups.

In other words, based on the MAHFP one might conclude that food security is comparable between the three groups, something which is not borne out by the HDDS. Without information on the true food security status of households, a 'gold standard', it is impossible to determine which of these two indicators is closer to the truth. It might even be that each indicator measures a completely different aspect of food security, or that the indicators do not measure food security at all. Before being able to draw conclusions on how well the HDDS measures food security, it should be established that the components of the indicator, the twelve food groups, behave in such a way that they can be logically added up into a single total score. This first step – the internal functioning of the indicator - is the subject of this study. We lack the 'gold standard' food security data required to take the second step – establishing whether the total score actually says something about food security. Thus, that the HDDS measures food access is assumed, not tested.

Methodology

Rasch models were developed by Rasch (1960) to measure an individual's level of a latent trait. The models assume that the probability of an individual's response to a question depends only on item difficulty and individual ability. In this study, the latent trait is assumed to be household food access, as suggested the developers of the HDDS (Swindale and Bilinsky, 2006). The food groups making up the indicator are the items. Rasch models do not depend on *a priori* assumptions about item difficulty. Rather, item difficulty is an outcome of the analysis. Rasch models are most frequently applied in education and psychology, but commonly used in other human sciences (Bond and Fox, 2001), and increasingly applied to medical research (Smith et al., 2008).

Rasch models have been used to study food security indicators before. They have been applied to test experience-based indicators, such as the core food security module (CFSM) developed by the US Department of Agriculture (Derrickson et al., 2000; Opsomer et al., 2003), Latin American Household Food Security Measurement Scale (ELCSA) (Toledo Vianna et al., 2012), Household Food Insecurity Access Scale (HFIAS) (Deitchler et al., 2010), and most recently, the Arab Family Food Security Scale (Sahyoun et al., 2014). Rasch models allow evaluating whether items are equally difficult in different cultural settings because

estimated item parameters are not sample specific (Casillas et al., 2006; Salzberger et al., 1999).

Rash analysis assumes hierarchical ordering of items. In the context of the HDDS, this implies that households consuming the most difficult item – i.e. the food group eaten only by those households with high food access - should also consume easier items. Although there is an extensive literature on dietary patterns which concludes that households shift to more expensive foods when their income increases (Thorne-Lyman et al., 2010), implying some hierarchy between foods, it is not clear to which extent this hierarchy is accurately captured by the food groups as defined in the HDDS. The hierarchical ordering of items is essential for the applicability of Rasch modelling, but cannot be tested directly. Not meeting this key assumption has several consequences, which are explained in the discussion section.

Two other conditions an indicator of food access should meet in order to be a valid and reliable proxy of the latent trait, household food access, could be tested directly by using Rasch analysis. First, the indicator needs to be robust to cultural differences. Hence, conditional on the latent trait, item difficulty should be consistent between countries, cultures, and food habits. Second, the probability of an affirmative response to an item (food group) needs to be stable over the latent trait, such that each food group contributes positively and significantly to the overall score on the indicator. These conditions are necessary for the indicator to reliably distinguish households with high food access from households with low food access and to allow cross-cultural and intertemporal comparison of households based on the HDDS.

Its most simple form, the 1PL Rasch model (equation 1), is based on the assumption that the probability of an affirmative answer to item i (e.g. consumption of a food group) by person p is determined by the difference between the person's ability θ_p (e.g. its food access status) and the difficulty of the item, β_i. In other words, the higher a person's food access status and the less 'difficult' a particular food group is, the more likely it is that this person is consuming that particular food group. Formally, the 1PL model is specified as follows:

$$\ln\left(\frac{P_{pi}}{1-P_{pi}}\right) = \theta_p - \beta_i \qquad (1)$$

This formula states that the log odds of the probability of an affirmative response of person p to item i is a linear function of the ability of person p (θ_p) and the difficulty of question i (β_i).

A poor item fit might indicate that the item does not measure the same latent trait as the other items, but it might also indicate that the item is not as strongly correlated with the latent trait as the other items. A simple 1PL Rasch model

assumes all food items are equally informative of a household's ability. The more flexible parameterization of the 2PL model allows testing the correlation of item i with the latent trait, by adding an interaction term, α_i:

$$\ln\left(\frac{P_{pi}}{1-P_{pi}}\right) = \alpha_i\theta_p - \beta_i \qquad (2)$$

The additional parameter, α_i, determines the discriminatory power of the items, i.e. it measures the extent to which an item helps to distinguish high from low performers. The larger is α_i, the more a small increase in θ increases the probability of an affirmative response to item i.

For interval scale measurement, each item should contribute positively to the latent trait, such that food access status increases with the consumption of each food group. In terms of the model, this implies $\alpha_i>0$. If α_i is not significantly different from zero, the probability of an affirmative response is no longer a function of θ. This implies that an individual with a highly diversified diet could not be distinguished from a household with a less diversified diet. More worrying are items (food groups) with a negative α_i. Such items showed an inverse relationship with the latent trait, implying that the probability of consuming food group i decreased with increasing food access. As the HDDS score equals the number of consumed food groups, food groups with an inverse relation with dietary diversity will bias HDDS downwards. Clearly, such items should not be included in a valid indicator.

A necessary pre-condition for any scale is that item response (food group consumption) should only depend on ability, not on any other individual- or household-specific characteristic. This pre-condition was checked using Differential Item Functioning (DIF) tests, which allows testing whether individuals with the same latent trait but different consumption preferences respond differently to items (Pallant and Tennant, 2007; Tennant and Conaghan, 2007). Consumption preferences are likely to differ between cultures and regions. For example, fish consumption might be common in coastal areas, but is linked to a highly diversified diet in rural areas. To verify this condition, prior knowledge of dietary patterns in the region was required.

For each subgroup of households in the sample, a refined indicator was constructed based on the relationship between individual items and overall score on the indicator. In a first step, food groups consumed by nearly all or none of the households were removed. Such items did not add value in distinguishing households with high food access from households with low food access. Furthermore, items with less than ten observations per binary choice alternative might cause estimates to become unstable (Linacre, 2002) and hence were removed. Second, the relationship of the remaining items with the total score on the indicator was visually verified with Item Response Functions (IRFs). Well-

functioning items should have a significant positive relationship with the overall score, indicating households consuming the food group had a higher probability of having higher food access. Badly functioning items were removed from the refined indicator. Item fit was further examined using item characteristic curves (ICCs), which show expected and observed probabilities for each item in a single graph (Bond and Fox, 2001).

The resulting refined indicators were tested for robustness and local independence. Robustness was checked by removing observations with low person-fit and verifying whether the ranking of items differed between the reduced sample and the full sample. Local independence requires that items are independent of one another, conditional on the value of the latent trait and was tested for by the significance of the correlation between response pairs (Ponocny, 2001; Tennant and Conaghan, 2007). All equations were estimated using Rasch analyses performed using *R version 2.12.1*, with packages *irtoys* and *eRm* (Mair and Hatzinger, 2007; Partchev et al., 2009).

Results

Consumption patterns of Colombian and Ecuadorian households were completely different, as is evident from Table 4.2 on page 98 and Table 4.3 on page 101. Hence, separate Rasch analyses were performed for each country. Differential Item Functioning showed the existence of distinct dietary patterns for Kichwa and immigrant households in Ecuador, requiring separate analyses for these two subgroups. Such a difference was not found in the Colombian sample. Therefore, three distinct analyses had to be performed, as shown in Figure 4.1. The consequently large amount of analyses performed implies that not all

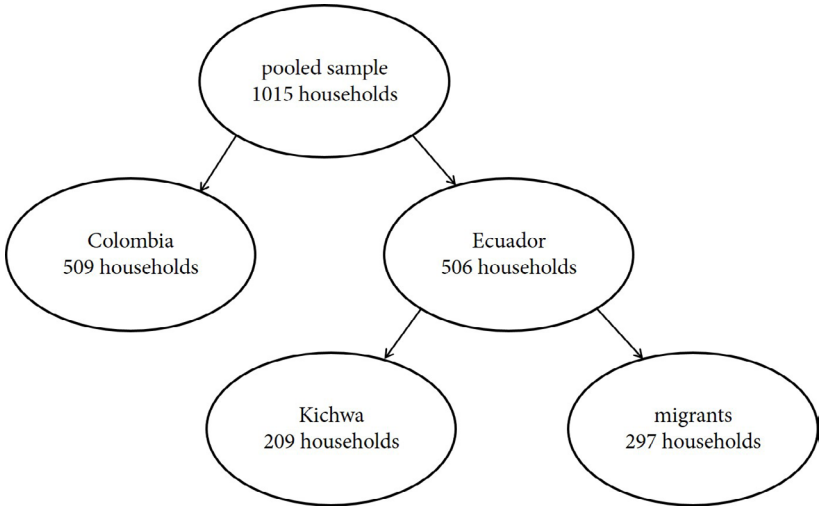

Figure 4.1. Structure of the results section

results could be reported in the main text; additional output is presented in the appendix.[5] In the next section (5.1), results of the analysis for Colombia will be discussed, followed by those for the DIF analysis in Ecuador (5.2) and the HDDS verifications for Kichwa (5.3) and migrant households (5.4).

Colombia

Food groups consumed by nearly all or very few households reduce the variation of the HDDS indicator and hence its efficiency. In the Colombian sample, this lack of variation was cause for concern: 99% of households consumed the food groups 1 (cereals), 2 (roots and tubers), 11 (sugar/honey) and 12 (other) during the 24 hours before the survey (Table 4.2). The nearly uniform consumption of these food groups meant they did not add explanatory power in differentiation between households with high and low food access. Therefore, their removal did not make the overall indicator less precise but was necessary to ensure stability of the estimates of the model (Linacre, 2002).

The relationship between individual items and the overall score was evaluated with Item Response Functions (IRFs) of an estimated 2PL Rasch model (Figure 4.2). IRFs showed the probability of an affirmative response for each item as a function of the latent trait, household food access. The higher was food access (on the horizontal axis), the higher should be the probability of consuming the food group(on the vertical axis). The numbers on the different curves correspond to the items (food groups) provided in Table 4.2. All food groups appeared to behave as expected: all curves show an upward slope.

Table 4.2. Food group consumption by Colombian households

#	Food group	By % of households
1	Cereals	99
2	Roots and tubers	99
3	Vegetables	49
4	Fruits	50
5	Meat	67
6	Eggs	66
7	Fish	6
8	Legumes	62
9	Milk/dairy	23
10	Oils/fat	86
11	Sugar/honey	99
12	Other	99

Note: food groups that were excluded from further analysis are underlined.

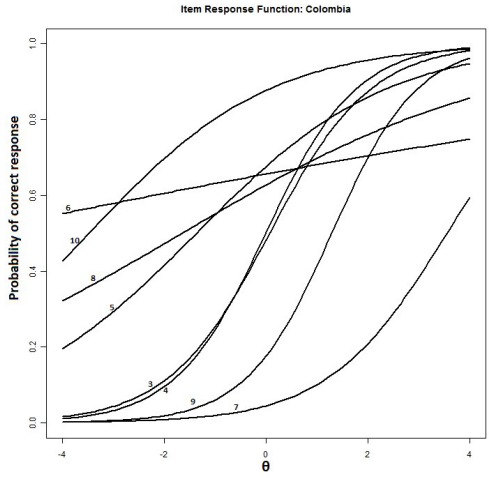

Figure 4.2. Item response functions (IRFs) Colombia

If two items had similar discriminatory power, α, but differed with respect to their difficulty, β, the curve of the most difficult item (higher β) would be plotted towards to the right-hand side of the figure. For instance, food group 3 (vegetables) and 7 (fish) had similar discriminatory power (α equalled 1.006 and 0.858 respectively), but vegetables (β = 0.07) was a considerably easier item than fish (β = 3.55). Hence, the IRFs of fish and vegetables were almost parallel, but the curve of vegetables was located to the left of the curve of fish.

The α's determine the slope of the IRFs: items with high discriminatory power have steeper slopes. For instance, food group 5 (meat) and food group 8 (legumes) had similar β's, but the slope of the IRF of meat was steeper than the slope of the IRF of legumes, because the latter had a smaller α. In other words, the food group meat had more power in differentiating between households with high and low food access.

The IRF of food group 6 (eggs) was rather flat, which indicated the probability of consuming eggs might be independent of the latent trait. A test confirmed that the discriminatory power of food group 6 was not significantly different from zero (p=0.22), so the item was removed from the refined scale. Eggs might not explain household food access because they are an important component of the daily diet in Colombia, independent of the socio-economic status of the household (Dufour et al., 1997). Most households might eat eggs frequently but not daily. In our sample, eggs were consumed by two-thirds of the interviewed households (Table 4.2). All seven remaining food groups had a positive and significant relationship with the latent trait, and were therefore included in the refined scale.

Item fit was verified by visual inspection of the Item Characteristic Curves (ICCs) for each of the seven remaining items. ICCs are similar to IRFs and show the

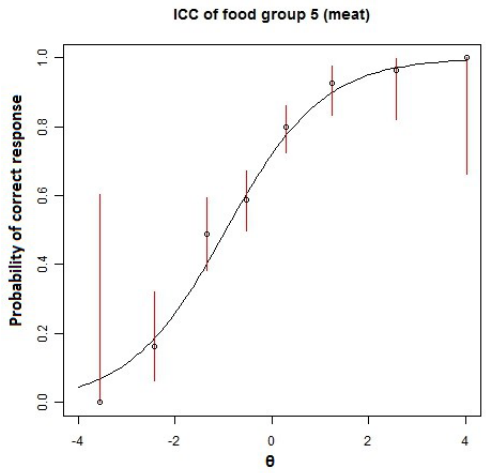

Figure 4.3. Item characteristics curve (ICC) of food group 5 (meat)

probability of consuming the food group (vertical axis) as a function of the household's food access (horizontal axis). ICCs also show the predicted probability of an affirmative response with its 95% confidence interval represented by vertical lines and the actual observed probability of an affirmative response represented by a dot. Item fit is high when predicted probabilities are close to expected probabilities. For example, for the food group meat (Figure 4.3), predicted probabilities corresponded well to actual observations. Results for other food groups were similar.

Results of the robustness check supported the model. Although removing the 6% of observations with low person-fit ($p<0.02$) did affect the size of the coefficients, it did not affect their difficulty rankings vis-à-vis one another. Local independence held. The nonparametric RM model test showed inter-item correlations between two out of 21 item-pairs, or roughly 10%. Based on the null hypothesis of independence this is no cause for serious concern. Further testing to find the source of dependencies based on principal component analysis resulted in a maximum eigenvalue of <1.3, with remaining eigenvalues slowly decreasing in size. Eigenvalues below 1.5 are generally considered to confirm local independence (Kahler and Strong, 2006).

Ecuador: Differential Item Functioning

The Amazon basin where the Ecuadorian data was collected had two ethnic groups with distinct dietary patterns. Originally the region was inhabited by the indigenous tribe of the Kichwa, but since the oil boom of the 1970s large groups of mestizo migrants have settled in the region and currently make up almost half the

Table 4.3. Food group consumption by Ecuadorian households

	Food group	% of Kichwa HHs (n=209)	% of migrant HHs (n=297)
1	Cereals	80	95
2	Roots and tubers	87	81
3	Vegetables	15	37
4	Fruits	26	40
5	Meat	52	66
6	Eggs	46	50
7	Fish	49	29
8	Legumes	18	56
9	Milk/dairy	7	27
10	Oils/fat	40	38
11	Sugar/honey	52	77
12	Other	54	86

Note: HHs = households

population (Lobao and Brown, 1998; Witt et al., 1999). A glance at the summary statistics for food groups consumption shows marked differences in diet between these groups (Table 4.3). Milk and dairy products were, for instance, consumed by only 7% of Kichwa households, while 27% of migrant households reported having consumed this food group in the previous day. This suggested that the pooling the data from Ecuador might cause validity problems.

A formal test confirmed the occurrence of Differential Item Functioning (Appendix F) between the ethnic groups ($P<0.001$), implying that a single index for the Ecuadorian case did not meet condition 3 of cultural robustness. When the items showing the strongest DIF were removed one by one until they no longer showed any DIF ($P=0.352$), only five food groups were left in the final model: 1, 3, 8, 9 and 11. Such a small number of groups is not very meaningful, as the resulting indicator can take only five values and is probably relatively insensitive to changes in food access. By not pooling the data, valuable within-group information on specific diets was preserved. Hence, the subsequent analysis was performed separately for each of the two cultural groups.[6]

Kichwa households

None of the food groups was consumed by so few or so many households to require removal from the indicator. The least frequently consumed food group, milk, was consumed by 15 households (Table 4.3). Item Response Functions for all food groups based on the 2PL model are shown in Figure 4.4. In order for

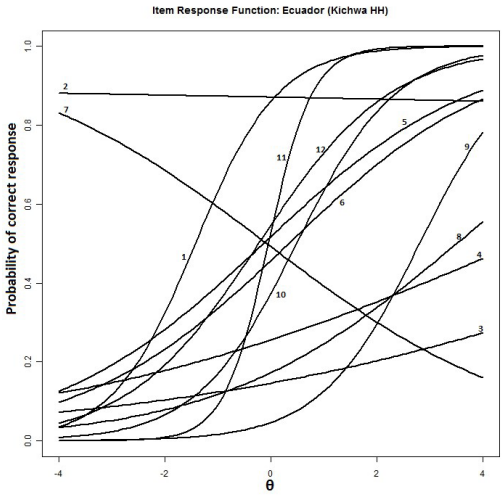

Figure 4.4. Item response functions Ecuador, Kichwa households

a food group to usefully contribute to the additive HDDS, the likelihood of its consumption needed to increase with an increase in the latent trait, reflected by a positive and significant slope. Food groups 2 (roots and tubers) and 7 (fish) both appeared to violate this condition.

The IRF of food group 2, roots and tubers, was a flat line. The item had low discriminatory power ($\alpha = 0.04$) and extremely low item difficulty ($\beta = -48.01$). The food group was consumed by 87% of Kichwa households, but their consumption was practically independent of their food access situation, meaning the group added no explanatory power to the overall indicator. It is likely that this food group was consumed by all households on a regular but not daily basis and therefore its consumption had no power in explaining household food access.

The negative slope on food group 7 (fish) indicated the likelihood of consuming fish decreased with increasing food access. The predicted likelihood of consuming fish decreased from 80% for households with little dietary diversity to less than 20% for households with a highly diversified diet. Previous research found fish to be an important part of the diet in Kichwa communities and consequently its consumption was common, although more so in rural communities than in towns (Webb et al., 2004). No sources were found mentioning an inverse relationship between income and fish consumption, although a possible explanation for the observed effect could be a development project of the provincial government of Napo which donated fish ponds to indigenous households in the region. Such a project was mentioned by respondents in a second survey round conducted in summer 2013.[7] If only food insecure households were eligible for this programme, it would explain the observed inverse relationship of fish consumption with overall dietary diversity.

Removing observations with low person-fit to verify robustness resulted in dropping 6% of observations. Differences between the full and reduced sample were negligible. Ranking was unaffected, and coefficient size hardly changed. Local independence tests based on inter-item correlation showed 6 out of 45 tested pairs, or 13%, showed significant correlation ($p<0.05$). Further analysis of the source of the variation indicated sampling variation rather than structural variation. The highest eigenvalue was 1.58; other eigenvalues were only slightly lower.

Migrant households

No food groups required removal from the refined indicator for migrant households because of too high or too low consumption frequency (Table 4.3 on page 101). The most frequently consumed food group was cereals, which was consumed by 95% of the population. Only 15 households did not report its consumption. Because this exceeded the critical threshold of ten observations per dichotomous choice alternative (Linacre, 2002), the food group was not removed.

Food groups 5 (meat) and 8 (legumes) appeared to have negative slopes (Figure 4.5), warranting their exclusion. Inspection of the coefficients of the 2PL model indeed showed that the slope of food groups 5 and 8 was negative ($\alpha = -0.11$ and $\alpha = -0.16$, respectively) and testing revealed that these slopes were not significantly different from zero at the 5% confidence level. These food groups were removed from the refined indicator.

Upon inspection of the ICC curves for migrant households, food group 7 (fish) was found to have low item fit. Many more households than predicted consumed fish at the lower tail of the distribution, meaning its consumption did not increase

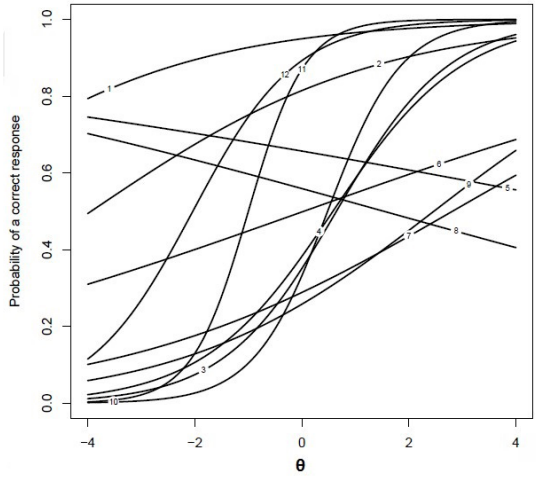

Figure 4.5. Item response functions Ecuador, migrant households

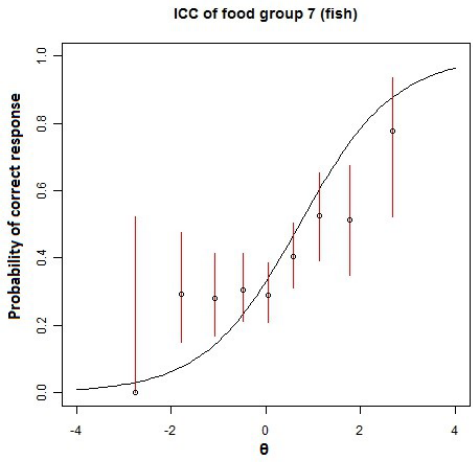

Figure 4.6. Item characteristics curve of food group 7 (fish) for migrant households

considerably with higher levels of food access (Figure 4.6). The Chi-squared test for item fit confirmed this conjecture; the null of good item fit was rejected ($p=0.013$). Therefore, food group 7 was removed from the refined scale. Re-testing showed the remaining items to have good fit.

The resulting scale was checked for robustness by removing the 5% of observations with low person-fit ($p<0.02$). This removal had a minimal effect on coefficient sizes and did not affect their ranking. Local independence did offer some cause for concern. Out of 36 item pairs tested for inter-item correlations, six were found to be significant (17%). Further testing of the source of the variation gave a maximum eigenvalue of 1.52. In other words, the observed local dependence was likely caused by sampling variation.

Cronbach's alpha

Cronbach's alpha is a widely used estimate of the internal consistency of the reliability of test scores. It is a function of the number of items in a test, the average covariance between items, and the variance of the total score. Alpha will increase if the covariance between items increases. Since this covariance is maximised when all items measure the same construct, the statistic is widely believed to indicate the degree to which a set of items measures the same latent variable. The statistic can take values between 0 and 1. As a rule of thumb, values above 0.9 indicate excellent internal consistency, values below 0.5 indicate unacceptable internal consistency (George and Mallery, 2003). Table 4.4 shows the results for some of the HDDS indicators used in this paper.

For none of the tested HDDS indicators the Cronbach's alpha exceeded 0.5, the

Table 4.4. Cronbach's alpha statistics for selected Household Dietary Diversity Score groups

Country	Sample	Cronbach's alpha	Confidence interval 5%	Confidence interval 95%
Ecuador	Full sample, n=515, items=12	0.42	0.33	0.50
	Kichwa, n=209, item=12	0.38	0.24	0.51
	Kichwa, n=209, item=10	0.49	0.37	0.67
	Migrants, n=297, items=12	0.30	0.18	0.43
	Migrants, n=297, items=9	0.44	0.33	0.55
Colombia	Full sample, n=509, items=12	0.35	0.26	0.44
	Full sample, n=509, items=7	0.39	0.29	0.48

minimum value for an indicator not to be considered inadequate. Removing food groups from the indicator did increase Cronbach's alpha: values for the refined indicator for each of the three groups exceeded values for the unrefined indicators. Although the alphas did increase, the improvement was insufficient to push them over the 0.5 threshold.

Discussion

In this paper the household dietary diversity score (HDDS) developed by the Food and Nutrition Technical Assistance (FANTA) project was analysed using Rasch models. In particular, it was verified whether the indicator met the criteria required for interval scale measurement. Meeting these criteria would imply the HDDS can be used as a household level indicator of food access. Such validity at household level is relevant for both development and research organisations, as it would allow attribution of project impact to specific outcomes. Rasch models allow differentiation between the discriminatory power and difficulty of items, revealing the relative importance of individual food groups in differentiating between levels of food access. In our data, this importance differed markedly between countries and between groups within Ecuador. Therefore, in its current form the HDDS did not meet the criteria.

For most application of Rasch modelling, the assumption of hierarchical ordering of the items is essential. In our application, this would imply that the food groups can be ranked ex-ante based on their difficulty. It also implies that a household that consumes the most difficult item should have consumed the other, easier, items. For dietary diversity, such a strict hierarchical ordering is difficult to establish, as it depends on locally prevailing market conditions (availability, price) and cultural preferences. Yet, in poor and food insecure regions, it is reasonable to assume that more food secure households consume more and less accessible food groups. It is hard to believe that households switch completely from one diet

to another or no longer consume certain food groups as they grow richer. In this sense, a hierarchical ordering of food groups is likely. On the other hand, changing dietary patterns may not completely be captured by the HDDS. For instance, food secure and food insecure households may both eat meat, but more food secure households may switch from chicken to beef. The HDDS is insensitive to such changes. A second argument supporting the applicability of Rasch analysis is the main conclusion - that some food groups are not correlated with the overall HDDS score or with the consumption of other food groups – could be reproduced using 'simple' descriptive and comparative statistics.[8] Therefore, even if the key assumption of hierarchical ordering was not met, robustness checks using similar statistics yield similar results. Importantly, in this paper Rasch analysis was not used to calibrate the indicator, which would require precise estimates of item parameters and hence would be more sensitive to the consequences of invalidating the assumption.

Assuming hierarchical ordering of the food groups, Rasch models can be used to test two conditions which a valid indicator of food access should meet: (i) cross-cultural validity and (ii) an increasing probability of an affirmative answer with increasing food access. The pooled data, combining data from Colombia and Ecuador, did not meet the condition of cross-cultural validity. More worryingly, even within the sample of Ecuadorian households, significant differences in dietary patters between Kichwa and migrant households were detected. Data had to be split into three different groups – Colombian, Kichwa, and migrant households – which were analysed separately. For each of these groups, items (food groups) not meeting the second condition were removed from the scale until a 'refined' HDDS was found that did meet all conditions. An overview of the three resulting scales is given in Table 4.5. It contains for each food group either the reason it was omitted from the scale or its difficulty ranking in the overall scale. The most difficult food groups were most likely to be eaten by households with the highest food access score.

There are large differences between the three refined indicators in the number of food groups they contain and in the difficulty ranking of these food groups. In the Colombian data, seven food groups made up the refined indicator: vegetables, fruits, meat, fish, legumes, diary, and oils. These results correspond well with the literature as the refined index mainly contains foods with high nutritional values such as fruits, vegetables, and animal source products. The results for the Ecuadorian subgroups were less convincing. For Kichwa households, the food groups roots and tubers, and fish were excluded from the final index and for migrant households the groups meat, fish, and legumes did not meet the conditions. Especially the non-inclusion of meat and fish in the overall index for both groups is cause for concern, as animal source foods are of crucial importance for macro and micro nutrient intake in developing countries (Murphy and Allen,

Table 4.5. Reason for removal and difficulty ranking of food groups for refined indicators

	Food group	Colombia (n=509)	Kichwa (n=209)	migrants (n=297)
1	Cereals	targeting	1	1
2	Roots and tubers	targeting	$\alpha = 0$	2
3	Vegetables	5	10	8
4	Fruits	4	9	7
5	Meat	3	3	$\alpha = 0$
6	Eggs	$\alpha = 0$	5	5
7	Fish	7	$\alpha < 0$	low item fit
8	Legumes	2	8	$\alpha = 0$
9	Milk/dairy	6	7	9
10	Oils/fat	1	6	6
11	Sugar/honey	targeting	4	4
12	Other	targeting	2	3

Note: numbers in columns indicate food group difficulty ranking (1 indicates the easiest food group); targeting indicates the food group was excluded because 99% of households consumed it; $\alpha = 0$ indicates the food group does not explain food access; $\alpha < 0$ indicates the food group has a negative relationship with food access, such that its consumption is associated with lower food access; low item fit indicates a significant difference between expected and predicted responses.

2003). Moreover, as there appears to be a direct link between consumption of animal source foods and dietary diversity (Brown et al. (2002), as cited in Ruel (2003)), the exclusion calls into question what the HDDS really measures.[9] When items are Giffen goods, such that their consumption decreases when their relative price falls, such as might be the case for fish in Ecuador, the imposed additive relationship between item consumption and food security no longer holds.

There were substantial differences in the importance of each food group in the overall index between countries and even within a country. This holds even though two culturally similar neighbouring countries were studied. In its current form, the HDDS has no cross-cultural validity, a problem previously mentioned but not tested by Ruel (2003). DIF-analysis showed that the indicator is not even necessarily valid within a country, as in Ecuador dietary patterns differed between groups with a different cultural background. This lack of cross-cultural validity is problematic as it prevents direct interpretation of the value of the overall indicator. Before interpreting this value, it is essential to have a thorough understanding of local dietary patterns, even when a survey or project concerns only a small area within a single country. Clearly, requiring extensive knowledge before being able to interpret a simple, easy-to-use indicator limits its usefulness for deployment in the rapid assessments required by development projects.

A potential cause of the limited accuracy of the HDDS at household level might be its focus on only the foods consumed in the last 24 hours before the survey

(Swindale and Ohri-Vachaspati, 2005). In that case, a straightforward way to overcome this inaccuracy is to increase the recall period. In a study using a 15 day recall period for dietary diversity, Drewnowski et al. (1997) noted diversity increased steeply over the first three days of recall, after which further increases became small. In other words, 24h recall might significantly underestimate true diversity when measuring dietary diversity at an individual or household level. Specifically, it might reduce the inaccuracy stemming from food groups that are eaten frequently, but not daily.

Other factors which might increase the construct validity of the indicator are re-defining the included food groups, adding weights, consumption frequency, and establishing minimum portion sizes. Food groups could be re-defined based on nutritional values, as is already being suggested specifically for iron deficiency (FAO, 2012) and is common in studies in the field of nutrition (Ruel, 2003). Weights could be added to account for the distinct nutritional value of food groups, as is already done by the Food Consumption Score used by the World Food Programme (WFP, 2008). The frequency of consumption might also be considered, which is particularly important in the presence of habit formation. Then, households might prefer those foods consumed as a child even when alternative food baskets become affordable (Atkin, 2013). Finally, minimum portion sizes should be considered. Ruel (2003) gives an example from Ghana, where fish consumption appeared high until it was found out fish meal was added in small amounts to porridge, obviously limiting its nutritional contribution. Different indicators take one or several of these factors into account, but knowledge of the contribution of each factor to the overall accuracy of the indicator is lacking. Further research is needed to specify and quantify the trade-offs involved.

Conclusion

The HDDS was developed as an easy-to-use and quick-to-implement survey-based assessment tool to allow measuring the impact on household food access of programs with improvements in food security as their core objective. Although its accompanying guidelines suggest that an "increase in household dietary diversity reflects an improvement in the household's diet", our results show the indicator should be cautiously interpreted. The HDDS does not allow comparing food access between different countries. Moreover, even in a small region within a single country, the indicator should not be used without sufficient knowledge of local dietary patterns. When dietary patterns differ between groups within a region, scores should not be aggregated for the region as a whole. Even within these relatively homogenous groups, there is a limited fit between included food groups and the underlying latent trait, such that the components of the indicator do not form a reliable way of measuring the variable of interest: food access.

Several problems were encountered with regard to the food groups making up the indicator. The gravest problem encountered was the inclusion of a food group with a negative relationship with the latent trait (Giffinity), implying that households were more likely to consume the food group when they had lower food access. Such items should never be included in an additive scale. In each of the three groups studied, there was at least one item which had no relationship with the latent trait, reducing the indicator's accuracy. Such items cause incorrect classification of households into food security states. Both problems might be avoided by re-defining the included food groups, adding weights, consumption frequency, and establishing minimum portion sizes. Until these issues are satisfactorily resolved, the HDDS should not be used as an indicator of the food access status of individual households.

Acknowledgements

We thank Paul de Boeck for his support with the methodology and interpretation of the results, and Carl Lachat for the insights regarding the nutritional implications.

Appendix A: Timing of the survey

The guidelines of the HDDS prescribe that in order for the indicator to "accurately capture changes in HDDS over time, data should be collected during the period of greatest food shortages (such as immediately prior to the harvest)" (Swindale and Bilinsky, 2006). The figures below show the share of households reporting to have insufficient food on a monthly basis, for each of the three groups. The survey

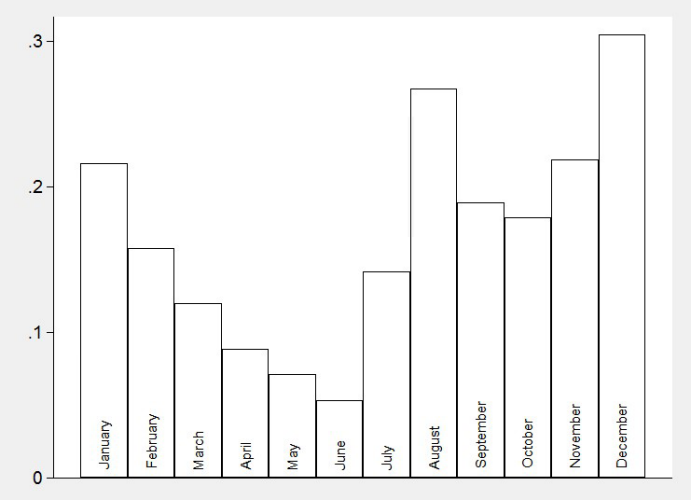

Figure 4.7. Seasonality in food insecurity for Colombian households

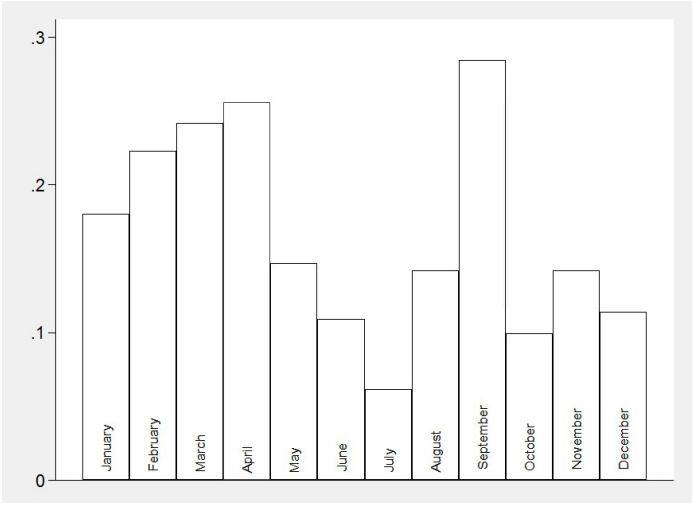

Figure 4.8. Seasonality in food insecurity for Kichwa households

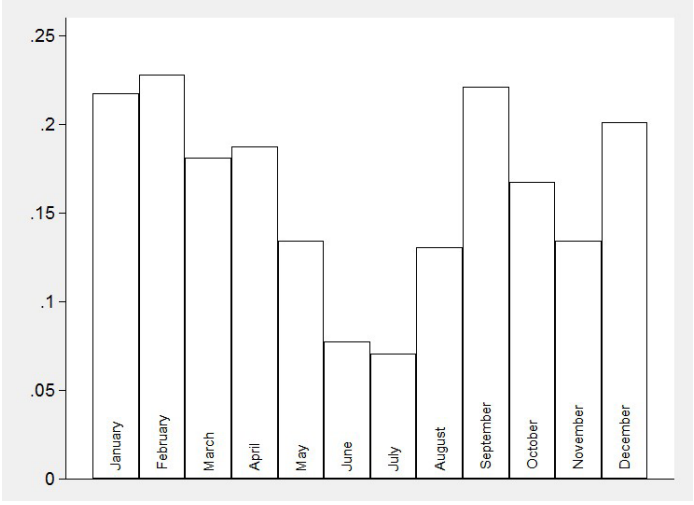

Figure 4.9. Seasonality in food insecurity for migrant households

was conducted in April-May. These months are among the most food secure for Colombian households (Figure 4.7) and relatively food insecure for Kichwa (Figure 4.8) and migrant households (Figure 4.9). It is interesting to note that it appears that for Colombian households the most food secure months are directly prior to the coffee harvest, which runs from May to August.

Appendix B: list of variables

MAHFP

Months of Adequate Household Food Provisioning, an indicator developed by FANTA which asks for each of the lasts 12 months whether the household had sufficient food. The value of the indicator is the number of months the household indicated not to have had sufficient food.

Ethnicity

The ethnic group the household head felt most closely related to. All interviewed households were ethnically homogenous. In Ecuador, the question was framed as a semi-closed multiple choice item with the following options (translated from Spanish): indigenous (if yes, which group), African-Ecuadorian, Mestizo, Mulato, White, Other (if yes, which), or No Response. All indigenous households were Kichwa, the original inhabitants of the zone. All migrant households were mestizo.

Family Size

The sum of the number of adults and children living in the same house and sharing the same food.

Dependency ratio

The number of children in the household divided by the number of adults in the household.

Household income

The sum of income from self-employment (agriculture, coffee, non-agriculture), wage (agriculture, non-agriculture), and income obtained from public and private transfers. Data were obtained using recall for the twelve individual months preceding the interview date. Reported incomes are the sum over these twelve months.

HDDS

The Household Dietary Diversity Score, developed by the FANTA project (Bilinsky and Swindale, 2010), consists of the twelve food groups shown in Appendix C. Respondents are asked which of these food groups they consumed in the past 24 hours, and their total score is the sum of their affirmative responses.

Appendix C: HDDS surveys

The used HDDS surveys were made more specific for each country by adding commonly consumed foods to the specification of the food groups.

Table 4.6. HDDS used in Colombia

26. Índice de diversidad dietética en el hogar (HDDS)

Ahora quisiera preguntarle sobre los tipos de alimentos que usted o cualquiera de los miembros de su familia comieron durante el día de ayer y en la noche.

LEA LA LISTA DE ALIMENTOS. SELECCIONE LA OPCIÓN "SÍ" DE LA CASILLA SI ALGÚN MIEMBRO DEL HOGAR CONSUMIÓ EL ALIMENTO NOMBRADO;

SELECCIONE LA OPCIÓN "NO" EN LA CASILLA SI NINGÚN MIEMBRO DEL HOGAR CONSUMIÓ EL ALIMENTO.

PREGUNTAS	CATEGORÍAS DE CLASIFICACIÓN
A. CEREALES ¿Algún cereal como el arroz, el maíz o el trigo, o algún producto elaborado con estos granos, como el pan, arepas, envueltos de choclo, fideos de trigo, hojaldres, tostadas, pasteles, o cualquier otro alimento hecho de mijo, sorgo, maíz, arroz, trigo, cebada, avena, etc.?	1. Sí ☐ 0. No ☐
B. RAÍCES, TUBÉRCULOS y PLÁTANOS ¿Papas, batata, yuca, arracacha, plátano, o cualquier otro alimento proveniente de raíces, tubérculos o plátanos?	1. Sí ☐ 0. No ☐
C. VERDURAS ¿Verduras?	1. Sí ☐ 0. No ☐
D. FRUTAS ¿Frutas?	1. Sí ☐ 0. No ☐
E. CARNE, POLLO, DESPOJOS ¿Carne de vaca, de cerdo, de cordero, de cabra, de conejo, de caza silvestre, cuy, pavo, pollo, pato u otras aves, hígado, riñón, corazón u otras carnes de órganos?	1. Sí ☐ 0. No ☐
F. HUEVOS ¿Huevos?	1. Sí ☐ 0. No ☐
G. PESCADO Y MARISCOS ¿Pescado o mariscos frescos o secos?	1. Sí ☐ 0. No ☐
H. LEGUMBRES/LEGUMINOSAS/FRUTOS SECOS ¿Alimentos a base de frijoles, arvejas, lentejas o frutos secos?	1. Sí ☐ 0. No ☐
I. LECHE Y PRODUCTOS LÁCTEOS ¿Queso, yogurt, leche u otros productos lácteos?	1. Sí ☐ 0. No ☐
J. ACEITES/GRASAS ¿Alimentos a base de aceite, grasa o mantequilla?	1. Sí ☐ 0. No ☐
K. ÁZUCAR/MIEL ¿Azúcar, miel o panela?	1. Sí ☐ 0. No ☐
L. ALIMENTOS DIVERSOS ¿Otros alimentos, como condimentos, café, té?	1. Sí ☐ 0. No ☐

Table 4.7. HDDS used in Ecuador

26. Índice de diversidad dietética en el hogar (HDDS)

Ahora quisiera preguntarle sobre los tipos de alimentos que usted o cualquiera de los miembros de su familia comieron durante el día de ayer y en la noche.

LEA LA LISTA DE ALIMENTOS. SELECCIONE LA OPCIÓN "SÍ" DE LA CASILLA SI ALGÚN MIEMBRO DEL HOGAR CONSUMIÓ EL ALIMENTO NOMBRADO;

SELECCIONE LA OPCIÓN "NO" EN LA CASILLA SI NINGÚN MIEMBRO DEL HOGAR CONSUMIÓ EL ALIMENTO.

PREGUNTAS	CATEGORÍAS DE CLASIFICACIÓN
A. CEREALES ¿Algún cereal como el arroz, el maíz o el trigo, o algún producto elaborado con estos granos, como el pan, la galleta, la humita, etc.?	1. Sí ☐ 0. No ☐
B. RAÍCES, TUBÉRCULOS y PLÁTANOS ¿Papas, camote, yuca, mandioca o cualquier otro alimento proveniente de raíces o tubérculos?	1. Sí ☐ 0. No ☐
C. VERDURAS ¿Verduras?	1. Sí ☐ 0. No ☐
D. FRUTAS ¿Frutas?	1. Sí ☐ 0. No ☐
E. CARNE, POLLO, DESPOJOS ¿Carne de vaca, de cerdo, de cordero, de cabra, de conejo, de caza silvestre, pollo, pato u otras aves, hígado, riñón, corazón u otras carnes de órganos?	1. Sí ☐ 0. No ☐
F. HUEVOS ¿Huevos?	1. Sí ☐ 0. No ☐
G. PESCADO Y MARISCOS ¿Pescado o mariscos frescos o secos?	1. Sí ☐ 0. No ☐
H. LEGUMBRES/LEGUMINOSAS/FRUTOS SECOS ¿Alimentos a base de frijoles, arvejas, lentejas o frutos secos?	1. Sí ☐ 0. No ☐
I. LECHE Y PRODUCTOS LÁCTEOS ¿Queso, yogurt, leche u otros productos lácteos?	1. Sí ☐ 0. No ☐
J. ACEITES/GRASAS ¿Alimentos a base de aceite, grasa o mantequilla?	1. Sí ☐ 0. No ☐
K. ÁZUCAR/MIEL ¿Azúcar o miel?	1. Sí ☐ 0. No ☐
L. ALIMENTOS DIVERSOS ¿Otros alimentos, como condimentos, café, té?	1. Sí ☐ 0. No ☐

Appendix D: output tables of 2PL models

Table 4.8. 2PL model Colombia including eggs

Food group	alpha	beta	se alpha	se beta	t-values alpha
Vegetables	1.005	0.069	0.25	0.108	4.023
Fruits	1.128	-0.007	0.287	0.099	3.937
Meat	0.538	-1.368	0.185	0.455	2.91
Eggs	0.11	-5.901	0.146	7.829	0.754
Fish	0.858	3.552	0.306	1.045	2.806
Legumes	0.316	-1.651	0.148	0.794	2.139
Milk/dairy	1.192	1.292	0.331	0.257	3.601
Oils/fat	0.564	-3.48	0.225	1.259	2.509

Table 4.9. 2 PL model Kichwa including roots/tubers and fish

Food group	alpha	beta	se alpha	se beta	t-values alpha
Cereals	1.281	-1.424	0.433	0.343	2.955
Roots and tubers	-0.022	84.830	0.293	1102.814	-0.077
Vegetables	0.197	8.949	0.286	12.821	0.689
Fruits	0.229	4.662	0.215	4.325	1.067
Meat	0.502	-0.141	0.215	0.297	2.337
Eggs	0.511	0.339	0.213	0.314	2.395
Fish	-0.406	-0.073	0.204	0.356	-1.995
Legumes	0.447	3.507	0.262	1.947	1.703
Milk/dairy	1.073	2.808	0.481	0.956	2.229
Oils/fat	1.066	0.485	0.297	0.186	3.594
Sugar/honey	2.410	-0.048	1.062	0.107	2.268
Other	0.809	-0.229	0.253	0.204	3.197

Table 4.10. 2 PL model migrants including meat and legumes

Food group	alpha	beta	se alpha	se beta	t-values alpha
Cereals	0.397	-7.390	0.348	6.211	1.140
Roots and tubers	0.376	-3.938	0.221	2.218	1.704
Vegetables	0.953	0.645	0.277	0.201	3.444
Fruits	0.825	0.580	0.243	0.211	3.392
Meat	-0.106	6.107	0.176	10.097	-0.605
Eggs	0.198	0.034	0.166	0.591	1.193
Fish	0.321	2.812	0.192	1.658	1.670
Legumes	-0.155	1.534	0.164	1.774	-0.945
Milk/dairy	0.429	2.465	0.201	1.112	2.137
Oils/fat	1.453	0.481	0.467	0.142	3.111
Sugar/honey	1.903	-0.999	0.629	0.183	3.026
Other	1.040	-2.037	0.321	0.484	3.243

Appendix E: In- and out-fit statistics

Table 4.11. In- and outfit statistics Colombia (final model)

Food group	1 PL	In-fit	Out-fit
Vegetables	-0.123	0.885	0.823
Fruits	-0.193	0.882	0.841
Meat	-0.950	0.981	0.968
Fish	3.001	0.799	1.040
Legumes	-0.752	1.044	1.058
Milk/diary	1.199	0.811	0.727
Oils/fat	-2.182	0.871	0.829

Table 4.12. In- and outfit statistics Kichwa (final model)

Food group	1 PL	Infit	Outfit
Cereals	-2.238	0.874	0.758
Vegetables	1.294	0.994	1.068
Fruits	0.550	0.994	1.048
Meat	-0.695	1.066	1.120
Eggs	-0.435	1.037	1.007
Legumes	1.035	0.965	0.919
Milk/diary	2.141	0.808	0.584
Oils/fat	-0.150	0.913	0.911
Sugar/honey	-0.695	0.814	0.746
Other	-0.805	0.963	0.960

Table 4.13. In- and outfit statistics migrants (final model)

Food group	1 PL	Infit	Outfit
Cereals	2.625	0.795	0.747
Roots and tubers	1.150	1.021	0.949
Vegetables	-0.989	0.953	0.910
Fruits	-0.880	0.950	0.960
Eggs	-0.428	1.122	1.158
Fish	-1.388	1.049	1.187
Milk/diary	-1.533	1.031	1.090
Oils/fat	-0.974	0.763	0.689
Sugar/honey	0.910	0.756	0.682
Other	1.507	0.825	0.839

Appendix F: Differential item functioning, Ecuador

Table 4.14. 1PL model with interaction between item and household groups (DIF)

Food group	Beta	se Beta	t-value	p-value
group (0=migrant HH; 1=Kichwa HH)	0.734	0.160	4.577	0.000
HDDS 1	0.748	0.302	2.475	0.013
HDDS 2	2.235	0.330	6.769	0.000
HDDS 3	-2.462	0.309	-7.981	0.000
HDDS 4	-1.440	0.253	-5.690	0.000
HDDS 5	-0.235	0.224	-1.045	0.296
HDDS 6	-0.252	0.224	-1.127	0.260
HDDS 7	0.412	0.225	1.830	0.067
HDDS 8	-2.493	0.286	-8.729	0.000
HDDS 9	-3.442	0.421	-8.170	0.000
HDDS 10	-0.402	0.228	-1.759	0.079
HDDS 11	-0.535	0.227	-2.359	0.018
HDDS 12	-0.667	0.232	-2.871	0.004
Group 1 (kichwa) *HDDS2	-0.977	0.205	-4.763	0.000
Group 1 (kichwa) *HDDS3	-0.095	0.197	-0.482	0.630
Group 1 (kichwa) *HDDS4	-0.401	0.188	-2.132	0.033
Group 1 (kichwa) *HDDS5	-0.429	0.184	-2.325	0.020
Group 1 (kichwa) *HDDS6	-0.653	0.183	-3.558	0.000
Group 1 (kichwa) *HDDS7	-1.180	0.185	-6.365	0.000
Group 1 (kichwa) *HDDS8	0.180	0.193	0.934	0.350
Group 1 (kichwa) *HDDS9	0.057	0.221	0.258	0.797
Group 1 (kichwa) *HDDS10	-0.777	0.185	-4.207	0.000
Group 1 (kichwa) *HDDS11	-0.131	0.187	-0.700	0.484
Group 1 (kichwa) *HDDS12	0.105	0.193	0.546	0.585

Table 4.15. ANOVA test for DIF

	Df	AIC	BIC	Log. lik	Deviance	Chisq	Df	Pr(>Chisq)
model1PL_gr	14	7067	7161	-3520	7039			
model1PL_gr_int	25	6892	7060	-3421	6842	197	11	<2e-16

Appendix G: Model tests comparing 2PL with 1PL

Refitting the final models with a 1PL model (Table 4.16, Table 4.17, and Table 4.18), shows that the 1PL models fits the data only slightly less accurately than the 2PL models (e.g. the BIC criteria favors 1PL over 2PL models, while AIC and LR test prefer 2PL models).

Table 4.16. 1PL versus 2PL Colombia

	AIC	BIC	log.Lik	LRT	df	p.value
model1PL	3872	3906	-1928	8		
model2PL	3866	3926	-1919	17.75	14	0.007

Table 4.17. 1PL versus 2PL Kichwa

	AIC	BIC	log.Lik	LRT	df	p.value
model1PL	2334	2371	-1156	11		
model2PL	2326	2393	-1143	25.48	20	0.002

Table 4.18. 1PL versus 2PL migrants

	AIC	BIC	log.Lik	LRT	df	p.value
model1PL	3253	3294	-1616		11	
model2PL	3229	3303	-1594	42.61	20	<0.001

References

Arimond, M., Ruel, M.T., 2004. Dietary diversity is associated with child nutritional status: Evidence from 11 demographic and health surveys. The Journal of Nutrition 134, 2579-2585.

Arimond, M., Wiesmann, D., Becquey, E., Carriquiry, A., Daniels, M.C., Deitchler, M., Fanou-Fogny, N., Joseph, M.L., Kennedy, G., Martin-Prével, Y., Torheim, L.E., 2010. Simple food group diversity indicators predict micronutrient adequacy of women's diets in 5 diverse, resource-poor settings. The Journal of Nutrition 140, 2059-2069.

Atkin, D., 2013. Trade, Tastes, and Nutrition in India. American Economic Review 103, 1629-1663.

Bennett, M., 1941. Wheat in national diets. Wheat Studies 18, 37-76.

Bilinsky, P., Swindale, A., 2010. Months of adequate household food provisioning (MAHFP) for measurement of household food access: Indicator guide (v.4.), Food

and nutrition technical assistance project. Academy for Educational Development, Washington D.C.

Bond, T.G., Fox, C.M., 2001. Applying the Rasch model. Fundamental measurement in the human sciences. Lawrence Erlbaum Associates, Inc., Mahwah, NJ.

Brown, K.H., Peerson, J.M., Kimmons, J.E., Hotz, C., 2002. Options for achieving adequate intake from home-prepared complementary foods in low-income countries, in: Black, R.E., Fleischer Michaelsen, K. (Eds.), Public health issues in infant and child nutrition. Nestlé Nutrition Institute, Philadelphia, PA.

Casillas, A., Schulz, E.M., Robbins, S.B., Santos, P.J., Lee, R.M., 2006. Exploring the Meaning of Motivation Across Cultures: IRT Analyses of the Goal Instability Scale. Journal of Career Assessment 14, 472-489.

DANE, 2011. Anexo pobreza según departamentos. Colombia: Departamento Administrativo Nacional de Estadística.

De Ayala, R.J., 2013. Theory and practice of item response theory. Guilford Publications.

Deitchler, M., Ballard, T., Swindale, A., Coates, J., 2010. Validation of a measure of household hunger for cross-cultural use. Food and Nutrition Technical Assistance II Project (FANTA-2), Washington, D.C.

Derrickson, J.P., Fisher, A.G., Anderson, J.E.L., 2000. The Core Food Security Module Scale Measure Is Valid and Reliable When Used with Asians and Pacific Islanders. The Journal of Nutrition 130, 2666-2674.

Drewnowski, A., Henderson, S.A., Driscoll, A., Rolls, B.J., 1997. The dietary variety score: Assessing diet quality in healthy young and older adults. Journal of the American Dietetic Association 97, 266-271.

Dufour, D.L., Staten, L.K., Reina, J.C., Spurr, G.B., 1997. Living on the edge: Dietary strategies of economically impoverished women in Cali, Colombia. American Journal of Physical Anthropology 102, 5-15.

Faber, M., Schwabe, C., Drimie, S., 2009. Dietary diversity in relation to other household food security indicators. International Journal of Food Safety, Nutrition and Public Health 2, 1-15.

FAO, 1996. Rome declaration on world food security, World Food Summit. Food and Agriculture Organisation, Rome.

FAO, 2012. Guidelines for Measuring Household and Individual Dietary Diversity, in: Kennedy, G., Ballard, T., Dop, M. (Eds.). Food and Agriculture Organisation of the United Nations, Rome.

George, D., Mallery, P., 2003. SPSS for Windows step by step: A simple guide and reference. Allyn & Bacon, Boston, US.

Hatloy, A., Hallund, J., Diarra, M.M., Oshaug, A., 1999. Food variety, socioeconomic status and nutritional status in urban and rural areas in Koutiala (Mali). Public

Health Nutrition 3, 57-65.

Heady, D., Ecker, O., 2013. Rethinking the measurement of food security: From first principles to best practice. Food Security 5, 327-343.

Hoddinott, J., Yohannes, Y., 2002. Dietary diversity as a food security indicator, FCND Discussion Paper. International Food Policy Research Institute, Washington, D.C.

INEC, 2006. Las condiciones de vida de los Ecuatorianos. Instituto Nacional de Estadistica y Censos, Quito, Ecuador.

Jensen, R.T., Miller, N.H., 2010. A revaled preference approach to measuring undernutrition and poverty using calorie shares, NBER Working Paper.

Jones, A.D., Ngure, F.M., Pelto, G., Young, S.L., 2013. What are we assessing when we measure food security? A compendium and review of current metrics. Advances in Nutrition 4, 481-506.

Kahler, C.W., Strong, D.R., 2006. A Rasch model analysis of DSM-IV alcohol abuse and dependence items in the National Epidemiological Survey on Alcohol and Related Conditions. Alcoholism: Clinical and Experimental Research 30, 1165-1175.

Kennedy, G., Berardo, A., Papavero, C., Horjus, P., Ballard, T., Dop, M., Delbaere, J., Brouwer, I.D., 2010. Proxy measures of household food consumption for food security assessment and surveillance: Comparison of the household dietary diversity and food consumption scores. Public Health Nutrition 13, 2010-2018.

Kennedy, G.L., Pedro, M.R., Seghieri, C., Nantel, G., Brouwer, I.D., 2007. Dietary diversity score is a useful indicator of micronutrient intake in non-breast-feeding Filipino children. The Journal of Nutrition 137, 472-477.

Leroy, J.L., Ruel, M., Frongillo, E.A., Harris, J., Ballard, T.J., 2015. Measuring the Food Access Dimension of Food Security: A Critical Review and Mapping of Indicators. Food and Nutrition Bulletin 36, 167-195.

Linacre, J.M., 2002. Understanding Rasch measurement: Optimizing rating scale category effectiveness. Journal of Applied Measurement 3, 85-106.

Lobao, L.M., Brown, L.A., 1998. Development context, regional differences among young women, and fertility: The Ecuadorean Amazon. Social Forces 76, 819-848.

Mair, P., Hatzinger, R., 2007. CML based estimation of extended Rasch models with the eRm package in R. Psychology Science 49, 26.

Maslow, A.H., 1943. A theory of human motivation. Psychological Review 50, 370-396.

Moursi, M.M., Arimond, M., Dewey, K.G., Trèche, S., Ruel, M.T., Delpeuch, F., 2008. Dietary diversity is a good predictor of the micronutrient density of the diet of 6- to 23-month-old children in Madagascar. The Journal of Nutrition 138, 2448-2453.

Murphy, S.P., Allen, L.H., 2003. Nutritional importance of animal source foods. Journal of Nutrition 133, 3932-3935.

Opsomer, J.D., Jensen, H.H., Pan, S., 2003. An Evaluation of the U.S. Department of Agriculture Food Security Measure with Generalized Linear Mixed Models. The Journal of Nutrition 133, 421-427.

Pallant, J.F., Tennant, A., 2007. An introduction to the Rasch measurement model: An example using the Hospital Anxiety and Depression Scale (HADS). British Journal of Clinical Psychology 46, 1-18.

Partchev, I., Partchev, M.I., Suggests, M., 2009. Package 'irtoys'.

Pinstrup-Andersen, P., 2009. Food security: definition and measurement. Food Security 1, 5-7.

Ponocny, I., 2001. Nonparametric goodness-of-fit tests for the rasch model. Psychometrika 66, 437-459.

Rasch, G., 1960. Probabilistic models for some intelligence and attainment tests. University of Chicago Press, Chicago.

Rose, D., Charlton, K.E., 2002. Quantitative Indicators from a Food Expenditure Survey Can Be Used to Target the Food Insecure in South Africa. The Journal of Nutrition 132, 3235-3242.

Ruel, M.T., 2003. Operationalizing dietary diversity: A review of measurement issues and research priorities. The Journal of Nutrition 133, 3911-3926.

Sahyoun, N.R., Nord, M., Sassine, A.J., Seyfert, K., Hwalla, N., Ghattas, H., 2014. Development and Validation of an Arab Family Food Security Scale. The Journal of Nutrition.

Salzberger, T., Sinkovics, R.R., Schlegelmilch, B.B., 1999. Data Equivalence in Cross-Cultural Research: A Comparison of Classical Test Theory and Latent Trait Theory Based Approaches. Australasian Marketing Journal (AMJ) 7, 23-38.

Smith, A., Rush, R., Fallowfield, L., Velikova, G., Sharpe, M., 2008. Rasch fit statistics and sample size considerations for polytomous data. BMC Medical Research Methodology 8, 33.

Steyn, N.P., Nel, J.H., Nantel, G., Kennedy, G., Labadorios, D., 2006. Food variety and dietary diversity scores in children: Are they good indicators of dietary adequacy? Public Health Nutrition 9, 644-650.

Swindale, A., Bilinsky, P., 2006. Household dietary diversity score (HDDS) for measurement of household food access: Indicator guide (v.2), Food and Nutrition Technical Assistance Project. Academy for Educational Development, Washington D.C.

Swindale, A., Ohri-Vachaspati, P., 2005. Measuring household food consumption: A technical guide. Food and Nutrition Technical Assistance (FANTA) Project, Academy for Educational Development (AED), Washington, D.C.

Tennant, A., Conaghan, P.G., 2007. The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? Arthritis Care & Research 57, 1358-1362.

Thorne-Lyman, A.L., Valpiani, N., Sun, K., Semba, R.D., Klotz, C.L., Kraemer, K., Akhter, N., de Pee, S., Moench-Pfanner, R., Sari, M., Bloem, M.W., 2010. Household Dietary Diversity and Food Expenditures Are Closely Linked in Rural Bangladesh, Increasing the Risk of Malnutrition Due to the Financial Crisis. The Journal of Nutrition 140, 182S-188S.

Toledo Vianna, R., Hromi-Fiedler, A., Segall-Correa, A., Pérez-Escamilla, R., 2012. Household food insecurity in small municipalities in Northeastern Brazil: a validation study. Food Security 4, 295-303.

Webb, J., Mainville, N., Mergler, D., Lucotte, M., Betancourt, O., Davidson, R., Cueva, E., Quizhpe, E., 2004. Mercury in fish-eating communities of the Andean Amazon, Napo River Valley, Ecuador. EcoHealth 1, 59-71.

Webb, P., Coates, J., Frongillo, E.A., Rogers, B.L., Swindale, A., Bilinsky, P., 2006. Measuring Household Food Insecurity: Why It's So Important and Yet So Difficult to Do. The Journal of Nutrition 136, 1404S-1408S.

WFP, 2008. Food consumption analysis: Calculation and use of the food consumption score in food security analysis. World Food Programme (WFP), Rome.

Witt, J.M., Kakabadse, Y., Ortiz, R., Maldonado, L., 1999. Zonas intangibles de la Amazonia Ecuatoriana. Ministerio de Medio Ambiente, Quito, Ecuador.

Notes

1 See Appendix A for a discussion on the recommendations for survey timing included in the HDDS guidelines.

2 All interviewed households were ethnically homogeneous.

3 En las últimas 24 horas, comía algún cereal como el arroz, el maíz o el trigo, o algún producto elaborado con estos granos, como el pan, la galleta, la humita, etc.?

4 En las últimas 24 horas, comía algún cereal como el arroz, el maíz o el trigo, o algún producto elaborado con estos granos, como el pan, arepas, envueltos de choclo, fideos de trigo, hojaldres, tostadas, pasteles, o cualquier otro alimento hecho de mijo, sorgo, maíz, arroz, trigo, cebada, avena, etc.?

5 Specifically, tables containing parameter estimates of the 2PL models are shown in appendix D and in- and out-fit statistics in appendix E.

6 These samples could be considered on the small side for 2PL Rasch

analysis, which might lead to biased estimates (De Ayala, 2009). However, they are not problematically small for the purpose of this paper, since we do not rely on precise estimates of alphas and betas to draw our conclusions. Furthermore, model tests show only small differences with 1PL models, for which a sample size of 100 is already considered informative (appendix G).

7 We were not able to identify the project. Respondents were most likely referring to the "Piscicultura Sostenible para la Amazonía" project executed by the Centro Lianas (www.centrolianas.org).

8 Results available from the authors upon request.

9 In this paper, we assumed HDDS measures household food access, following the stated intention of its developers (Swindale and Bilinsky, 2006). Evidence supporting this claim is limited, warranting further research.

Chapter 5

The effect of specialty coffee certification on household livelihood strategies

based on

Vellema, W., Buritica Casanova, A., Gonzalez, C., D'Haese, M., 2015. *The effect of specialty coffee certification on household livelihood strategies and specialisation*. Food Policy 57, 13-25.

Abstract

Farm certification is rapidly becoming a sine qua non for access to specialty coffee markets. At the centre of these changes are smallholder coffee producers, responsible for 80% of global coffee production. Although rural households are known to depend on more than agricultural production alone, the literature on specialty coffee and certification has rather narrowly focused on coffee income and production. In this study, broader impacts on livelihood strategies are explicitly taken into account. Household income was decomposed into categories corresponding to specific income-generating activities and coffee income was broken down into price, yield and area effects. Results show that coffee certification encourages farmers to specialize in coffee production, increasing coffee income but not total household income, at least not in the short run. The time and effort required to attain the higher coffee income offered by certified production means farmers have to give up other activities. This substitution effect reduces out the effect on total income, which might suggest the additional labour effort required for certified coffee production reduces time spent on other activities, reducing income from these sources.

Introduction

Specialty coffee is considered to include all coffees that are differentiated, which means they are able to earn a premium (ICO et al., 2000; Ponte, 2002). Daviron and Ponte (2005) distinguish three types of differentiation: based on in-person service, material, or symbolic attributes. Only differentiation by material and symbolic attributes affect producers, as in-person service value can only be created at the point of consumption through interaction between provider and consumer and among consumers. Material attributes in the case of coffee relate to its intrinsic quality, whereas symbolic attributes are created through indications of geographical origin or sustainability labels. This latter group - coffee differentiated based on any kind of certification or label - is often referred to as sustainable coffee (Kolk, 2013; Raynolds, 2009).

Differentiation of coffee based on symbolic attributes started in the late 1970s and 1980s, when civil society movements started challenging the image of coffee as a tropical commodity by starting to sell fair trade and organic coffees out of alternative trade shops. To maintain differentiation and allow entry into mainstream markets, their ideals on social and environmental issues were formalized into labels and corresponding certifications (Raynolds, 2000). This formalization process coincided with increased popular consumption of coffee beverages differentiated by preparation method and often distributed through dedicated sales outlets, the so-called 'Latte revolution' (Ponte, 2002). More recently, single origin coffees are entering the market, often sold through specialty coffee houses.

Few statistics are available on the market share of specialty coffee, but those that do exist indicate its share is large and growing rapidly. In 2012, specialty coffee represented 37% of sales by volume and 50% by value in the United States (SCAA, 2012). Globally, sustainable coffee sales in terms of volume increased by 433% in the period 2004-09 (Potts et al., 2010). At current growth rates, sustainable coffee is expected to grow to 20-25% of the market by 2015 (ITC, 2011). This trend is likely to accelerate with the increasing participation of big retailers (Daviron and Vagneron, 2011; Elder et al., 2014).

Worldwide an estimated 26 million people are employed by the coffee sector, many of whom are smallholder farmers depending primarily on coffee to make a living (ICC, 2010; UNCTAD, 2003). The emergent pervasiveness of sustainable coffee has led to numerous studies comparing certified with non-certified chains and comparing different certifications in terms of (farm-gate) price, yield, profits, or production costs (Barham and Weber, 2012; Valkila, 2009; Weber, 2011) or analysing the impact of certification on outcome indicators other than income, such as education (Méndez et al., 2010; Ruben and Fort, 2012; Ruben et al., 2009) and vulnerability (Bacon, 2005; Bacon et al., 2014). These studies are characterized

by their focus on coffee production and income, and their analytical dependence on mean comparison, often enriched with qualitative evidence.

Although the relationship between certifications and coffee income is of interest in itself, most coffee producers cannot survive from coffee production alone (Bacon et al., 2008; Forero Álvarez and Furio, 2010). Rural households across the world depend on a broad set of income-generating activities to sustain their livelihoods (Davis et al., 2010; Ellis, 1998, 2000; Reardon et al., 2001). This importance of income apart from coffee production is corroborated by Barham et al. (2011), who show that even relatively specialised coffee farmers selling into specialty coffee markets obtain higher returns from diversification into labour markets. Recent work conducted in Nicaragua shows limited investment in capital for coffee production, especially among small-scale producers with diversified livelihoods (Donovan and Poole, 2014). If access to specialty markets allows producers to diversify into lucrative income-generating activities outside of coffee production, a focus on its relationship with coffee income alone will underestimate its impact.

In this paper we will take a first step towards measuring the relationship between access to specialty markets and livelihood strategies, based on a group of small-scale coffee producers in the south of Colombia that is strongly affected by changes in the coffee market. To show the extent of income diversification into activities beyond coffee production, current income portfolios of these coffee producers will be discussed in some detail. Subsequently, based on factors derived from the livelihood literature, barriers and opportunities driving observed income diversification strategies of coffee producers are identified, focusing on the effect of farm certification. Overall income effects and impact pathways are also considered. In the implications section, our findings are placed in the broader context of agricultural development and poverty alleviation and implications for policy makers are discussed.

Background

Coffee in Colombia

Coffee is grown throughout the hilly and mountainous areas of Colombia by over 563,000 mostly small-scale coffee producers. Colombia is one of the largest coffee producers in the world and uniquely positioned to benefit from the increasing demand for specialty coffee. In 2012, the year of our study, it ranked as the fourth largest overall producer and the second largest in Arabica coffee (ITC, 2012). For Colombian milds, the highest value category defined by the international coffee organisation (ICO), it is the largest producer. The high quality of Colombian coffee can be attributed to its high agroecological diversity, its location close to the equator, and the consistent institutional support for coffee farmers.

The Colombian Ministry of Agriculture and Rural Development (MARD) has several active programs on agricultural extension and income support of coffee producers. Agricultural extension consists of credit support and training in agronomy and farm management. There is a dedicated policy to help producers take the steps required to have their coffee qualified as specialty, to capture associated price premiums.[1] The income support program was re-branded Protección al Ingreso Cafetero (PIC) after strikes of coffee producers during the price slump in early 2013. When the domestic coffee price, as published daily by the national federation of coffee producers (FNC), is below the minimum price specified in the program at the moment of sale, producers are eligible for a fixed subsidy pay-out per kilo of coffee sold. All these support programs are implemented by the FNC, a semi-governmental organisation founded in 1927 which represents all coffee producers.[2] FNC buying offices are densely scattered throughout Colombia's coffee regions, providing a guaranteed market for all coffee offered for sale.

The department of Nariño, in the southwest of Colombia, is particularly well-suited for growth of high quality specialty coffee. Although only 3% of Colombia's coffee is produced here, it has the potential to produce coffee of exceptional quality because of its unique agro-ecological conditions, allowing coffee to be grown at extreme altitudes (Oberthür et al., 2011). In the 2010 Colombian Cup of Excellence competition, producers from this department claimed the first six spots and eight of the top ten[3]. Nariño is one of the fastest growing coffee regions in Colombia, with hectares planted equalling or surpassing those in the traditional Colombian coffee lands around Medellin (Rueda and Lambin, 2013). There are over 30,000 coffee producers in the region (FNC, 2010), most of whom are poor smallholders producing coffee for the commodity market (Garcia, 2003).

Access to specialty markets

When considering the group of specialty coffees differentiated based on sustainability labels, a distinction can be made between third-party labels and private labels. Third party labels such as Fair Trade, Organic, Bird Friendly and Rainforest Alliance are mostly mission-driven (Raynolds et al., 2007). They are also considered more legitimate (Gereffi et al., 2001). Private labels such as Starbucks' C.A.F.E. practices and Nestlé's Nespresso AAA certification are primarily aimed at ensuring a minimum quality of their coffee supply (Elder et al., 2014). The line between mission-driven and quality-driven standards is blurring[4] (Kolk, 2013; Raynolds et al., 2007; Raynolds, 2009), which results in a corresponding blurring of the distinction between coffee differentiated based on symbolic and material attributes. Companies are increasingly using labels to manage quality (Elder et al., 2014). The future of labels appears to be in increasing material quality. Price premiums unrelated to intrinsic quality are expected to dissipate as labels become

mainstream (de Janvry et al., 2010). Coffee certifications are now so common that they are becoming an entry requirement for access to upscale differentiated markets (Daviron and Vagneron, 2011; Elder et al., 2014; Raynolds, 2009).

On the producer level, it is important to distinguish between farm certification and certified sales. Not all coffee produced at certified farms receives a premium price, meaning far more certified coffee is produced than sold (Kolk, 2013). This situation is also prevalent in the field in Colombia (Table 5.1). Just over half of the interviewed households held farm certification. Having farm certification by no means guaranteed receiving price premiums: only 57% of certified farmers sold any coffee into certified marketing channels, and even fewer farmers (37%) sold all their coffee into specialty markets. Which and how many farmers are able to sell coffee into certified channels in any given year depends on demand, global supply of certified coffee, and quality, factors which are only partially under farmer control. Although having farm certification was no guarantee for certified sales (evidenced by the large number of certified farms which could not or only to some extent sell in certified markets), it certainly served as a prerequisite for access: only 3 farmers without farm certification (1%) were able to sell coffee into specialty markets.

Although participation into certification schemes is voluntary in principle, in practice it is often the buyers who select farmers. This was certainly the case for the most prevalent certifications in our study region, C.A.F.E. Practices and Nespresso AAA.[5] Starbucks started with C.A.F.E. Practices in 2004 (Starbucks, 2004, 2014) and told all their suppliers preference would be given to those involved in the program. Within a year 25% of suppliers had signed up (Alvarez et al., 2010). Empresas de Nariño, the local Starbucks supplier, was among the first to sign up, starting verification of farmers in 2004. The Nespresso AAA certification was developed in 2003 and started cooperating with the FNC in 2005. Nespresso buys all its Colombian coffee through the FNC. The local chapters in the departments of Nariño and Cauca were the first to sign up (Alvarez et al., 2010). Adoption was fast: the number of participating farmers in the departments increased from 500 in 2006 to 28,000 in 2013 (Nespresso, 2014).

The rapid growth in demand for specialty coffee has led buyers to rapidly increase the number of certified farmers. In Nariño, this has led buyers to go into more

Table 5.1. Certified sales versus certified production

Certified sales	Certified production		No certification	
None	113	43%	244	99%
Some	52	20%	1	0%
Only	98	37%	2	1%
Total	263	100%	247	100%

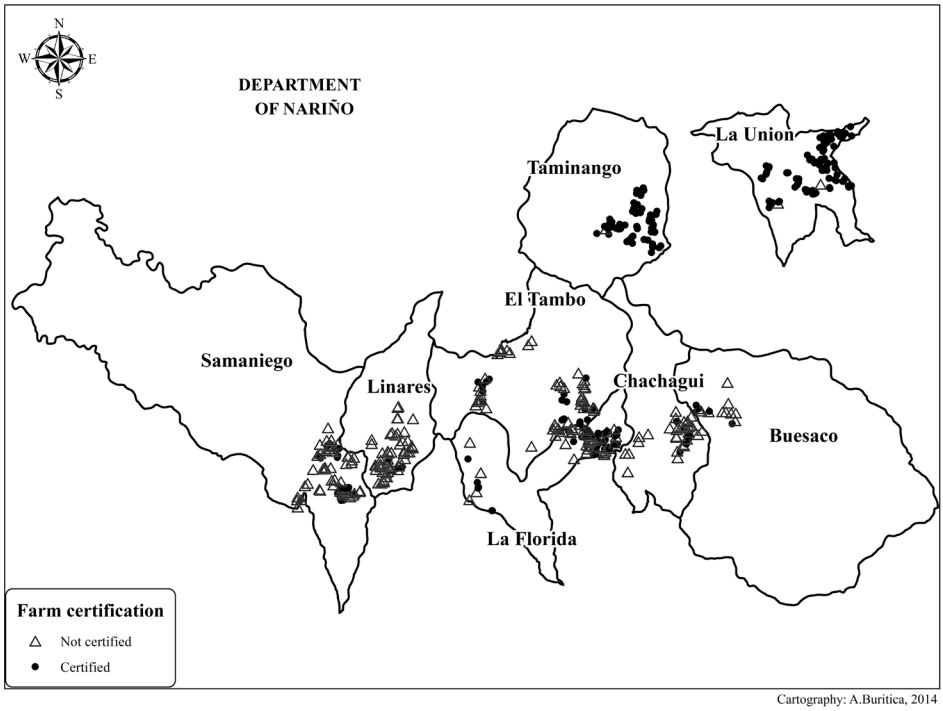

Figure 5.1. Location of certified and non-certified farms

remote coffee producing areas in search of more farmers, a process made possible by the successful reduction in guerrilla activity in the region (Ávila et al., 2014; Vargas and Caruso, 2014). At the start of 2012, when interviews were conducted, the various guerrilla groups active in Nariño had been pushed back to remote rural areas.

This pattern is clearly visible in the prevalence of certified farms (Figure 5.1 and Table 5.2). In the northern municipalities of La Unión and Taminango, almost

Table 5.2. Farm certification and size by municipality

Municipality	% certified	Farm (ha)	Coffee (ha)
La Union	98%	2.4	0.7
Taminango	99%	2.5	0.7
La Florida	53%	1.5	0.7
El Tambo	26%	2.5	0.4
Buesaco	33%	3.1	0.5
Chachagui	24%	2.5	0.5
Linares	5%	1.6	0.6
Samaniego	25%	1.2	0.9
Average	52%	2.0	0.6

all farms have been certified, whereas in the southern municipalities of Linares and Samaniego less than a quarter of farms hold certification. Certification rates in the central region are between these extremes. However, there is no clear observed pattern in farm size or specialization in coffee production between regions. Differences in certification rates between regions are not explained by easily observable farm characteristics. Rather, they appear to be driven primarily by the region in which farms are located.

Data

The department of Nariño can be divided into three distinct geological zones: the pacific coast, the Andes, and the Amazon basin. The Andean zone is the most densely populated – this is where coffee is grown (De la Hoz, 2007). In eight of the municipalities with the highest concentration of coffee production – La Union, Taminango, La Florida, El Tambo, Chachaguï, Buesaco, Linares, Samaniego - interviews were held with a total of 510 coffee-producing households in April-May 2012[6] (Figure 5.2). Out of 510 interviewees, 165 were female (32%) and 345 male (68%). These producers were selected through stratified random sampling from a list of coffee producers made available by the FNC, with stratification based on the number of coffee producers per municipality. In Colombia, every coffee producer is eligible for membership of the FNC. Although registration is not mandatory, practically all producers are members, as all governmental benefits including purchasing guarantees, minimum prices, income support and subsidized inputs and credit are distributed through the FNC. Most interviews were held with the head of the household. When the head of the household was not available, his or her closest relative was interviewed.

Interviews were conducted by trained local enumerators using a structured questionnaire developed by the Centro International de Agricultura Tropical (CIAT), a CGIAR centre based in Cali, Colombia. Based on this questionnaire, a detailed protocol was developed which was used to train the enumerators during a two-week training course including field trials. Surveys contained questions on household composition, farm characteristics, household income, and access to financial services. For each household member, age, years of education, and main occupation were recorded. Collected information on farm characteristics included land ownership and usage, as well as production, prices, income from farm sales and home consumption. Emphasis was placed on coffee production and commercialization. Raw data was cleaned and operationalized by the corresponding author, in cooperation with Martha Del Rio and Alexander Buritica Casanova from CIAT. Upon processing the information, production for home consumption was valued at farm gate prices and added to on-farm income. Income from other sources, including labour income, pensions, transfers, and

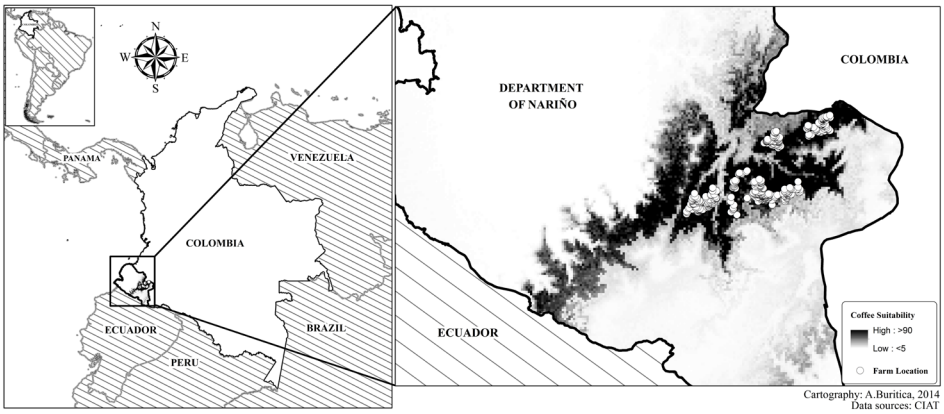

Figure 5.2. Study area: the Department of Nariño, Colombia

subsidies was recorded, and cross-verified with the stated occupation of household members. The final part of the survey concerned access to and use of financial services.

Interviews were conducted at the farm house, which allowed enumerators to register the GPS coordinates of each homestead. These coordinates were used to create the variable distance to nearest city. Distances from households to cities with over 10,000 inhabitants were calculated using cost distance algorithms, taking into account land cover, elevation, and the transportation network with assigned velocities (Elvidge et al., 1997; Small, 2004).

Income diversification

Rural households obtain income from a wide variety of sources: diversification is the norm (Davis et al., 2010; Ellis, 1998). In this regard, coffee producers in Nariño are no exception. Table 5.3 shows their income distribution over different sources, grouped into income from self-employment, wages, and transfers. Self-employment included income from farming, split into income from coffee production and farm income from other crops and animals, and income from self-employment outside of the agricultural sector. Wage income was grouped into agricultural and non-agricultural income, where agricultural income was primarily obtained from manual labour, often paid per day ('jornaleros'). Non-agricultural wage income included formal jobs for government or the private sector as well as informal jobs such as driving a bus. Transfer income was subdivided based on whether it derived from government (public) or private sources. The most important source of public transfers in our sample were pensions.

The first column in the table shows the relative importance of each income source for the sample as a whole. Coffee was by far the largest source of income

Table 5.3. Income distribution between and within income sources

Income source	% of total income	Use %	GINI	Corr	Avg. Income Share	SD of Avg. Inc. Share	Var %
Self-employed							
Agriculture	20	95	0.61	0.40	28	0.24	0.12
Coffee	46	92	0.62	0.80	44	0.29	0.51
Non-Agriculture	4	11	0.52	0.40	27	0.15	0.06
Wage							
Agriculture	11	51	0.44	0.30	34	0.22	0.05
Non-Agriculture	12	19	0.55	0.70	36	0.23	0.23
Transfer							
Public	5	47	0.64	0.27	12	0.16	0.05
Private	0	1	0.43	-0.05	38	0.31	0.00
Total income	100	100	0.54	1.00	100		1.00

in the sample (46%), which was not surprising given that our sample consisted of households identifying themselves as coffee producers ('caficultores'). What was surprising was the high share of agricultural versus non-agricultural income: a total of 77% of all income was derived from agriculture, most of it generated at the own farm. These figures indicate a relatively high dependency on farming income. Earlier studies on rural Colombia found off-farm income to contribute far more substantially. Deininger and Olinto (2001) found farm profits to contribute only 56% to total income, while the contribution of agricultural and non-agricultural wages (30%) and nonfarm enterprise profits and non-earned income (12.5%) were higher. Echeverri (1999) found off-farm income to contribute as much as 50%. Because both studies were based on a representative sample of rural households in Colombia, the relative importance of on-farm income in our sample might be typical for coffee producers, typical for Nariño, or typical for coffee producers in Nariño.

Not all households obtained income from all sources, nor did they obtain equal amounts from each source. The distribution of income per income source over households is shown by Gini coefficients (column 3). High Gini coefficients indicate the concentration of income from the respective source among a small part of the population. Almost all households obtained on-farm income from either coffee[7] or other agricultural production (column 2), but this income was distributed unevenly across the population. In contrast, only half of interviewed households received income from agricultural wage income, but this income was distributed far more evenly. The main explanation for this difference was the low remuneration of agricultural wage labour, evidenced by its small correlation

with income (column 4). This does not mean it was not an important source of income; those households engaging in off-farm agricultural wage labour on average received 34% of their income from the activity (column 5). However, non-agricultural wage labour was far more attractive. Although it contributed a similar share to total income for those active in it, 36% on average, its high correlation with total income shows how much better off these households were. The better remuneration of non-agricultural compared to agricultural wage labour is commonly encountered in the literature (Lanjouw and Shariff, 2002; Reardon, 1997; Reardon et al., 2001).

The final column shows total income variation originating from each source. Coffee is by far the biggest contributor to total income variation, due to the large share of coffee income in total income and its unequal distribution between households. The second biggest share of total income variation is explained by non-agricultural wage labour. Although only 19% of households engage in this activity, it explains 23% of total income variation.

The differences in income-earning portfolios between income strata is presented in Table 5.4, which shows the average income share of different income-earning activities per income quintile. Farmers with higher incomes obtain a larger share of their income from coffee production, with the share increasing from 33.0% for the lowest income quintile to 52.1% for the highest income quintile. The share of farming income remains stable over income groups; the increasing importance of coffee comes at the cost of crop and livestock income. This stable dependency on farming is not typical. Generally, the contribution of farming falls with income (Reardon et al., 2001). Dependency on the agricultural sector is higher for the first quintiles, but slowly falls with income. Agricultural wages contribute a particularly large share of income for the second (26.1%) and third (22.1%) quintile. The dependency of low income households on agricultural wages might be an indication of push diversification, born of necessity rather than preference (Barrett et al., 2001; Ellis, 2000). Employment outside of the agricultural sector is far more important for higher income groups. Although it contributes only

Table 5.4. Income by source and quintile

Income	Self-employed			Wage		Transfer	
quintile	Agri	Coffee	Non-agri	Agri	Non-agri	Public	Private
I	35.4	33.0	0.0	19.1	1.6	9.4	1.6
II	27.1	32.6	3.3	26.1	4.0	6.4	0.5
III	26.5	39.2	2.9	22.1	4.6	4.1	0.5
IV	23.4	45.1	3.9	12.9	11.5	3.2	0.0
V	18.3	52.1	4.6	6.3	13.0	5.7	0.0
Total	26.1	40.4	2.9	17.3	6.9	5.8	0.5

1.6% to total income in the first quintile, it contributes 13.0% for the fifth quintile. Such increasing importance of non-agricultural wages for higher income groups is also common in other parts of the world (Barrett et al., 2001; Davis et al., 2010). The shares of income obtained from self-employment outside of agricultural production and private transfers is low and do not differ much across income groups. Public transfers are more important for lower income groups.

Econometric approach

To assess whether and how having access to specialty markets influenced household diversification strategies, a distinction was made between the effect on the decision to undertake a certain activity and the effect on the amount of income derived from each activity. Both the participation decision and the income decision depend directly on the return to labour in each activity. Labour will be allocated to those activities with the highest return, taking into account decreasing returns to labour in each activity. Imperfect markets create differences between the market wage and shadow wage. The size of this difference depends on household-specific transaction costs and household asset endowments (de Janvry et al., 1991; de Janvry and Sadoulet, 2006; Dercon and Krishnan, 1996; Sadoulet et al., 1998). In order to accurately estimate the effect of access to specialty markets, factors influencing the shadow wage were controlled for. The estimated regression of the participation decision had the following form:

$$d_{ij} = \beta_{0j} + \beta_{1j}C_i + \beta_{2j}X_i + \beta_{3j}Z_i + \varepsilon_{ij} \quad (1)$$

Where dij is a binary variable indicating that household i participated in activity j. Activities were defined as belonging to one of five groups: on-farm agricultural production other than coffee, on-farm coffee production, self-employment outside of agriculture, and wage income from agricultural and non-agricultural activities. The estimated regression for the amount of income obtained from each activity was similar, differing only in its dependent variable:

$$y_{ij} = \beta_{0j} + \beta_{1j}C_i + \beta_{2j}X_i + \beta_{3j}Z_i + \varepsilon_{ij} \quad (2)$$

Where y_{ij} is the amount of income obtained by household i from activity j. The main variable of interest in both equations is C_i, which is a binary variable equal to one for households with farm certification, which equalled access to specialty markets in the study region. The certification is a prerequisite for access, but not a guarantee: of those households with certification – access – almost half did not have any certified sales (Table 1). X_i and Z_i are vectors of control variables known to influence the labour allocation decision through affecting the shadow wage. The vector X_i included important assets of rural households: land and labour[8], their relative size - measured by the land/labour ratio - and level of education.

To measure the importance of credit constraints, the amount of transfer income was included. The variable Z_i contains the distance to the nearest city to control for market access costs (Lanjouw et al., 2001). Details of variable construction and descriptive statistics are provided in Appendix B. Finally, α_j is an activity-specific intercept and ε_{ij} a cross-equation error term with a multivariate normal distribution, allowing error terms to correlate between activities.

The estimate for farm certification, our variable of interest, reflects the difference in income diversification of two households with the same observed characteristics in X_i and Z_i, but with one of those households having farm certification. This is not necessarily identical to a household's expected income diversification when it is selected to participate in certification and accepts. In the first case, the unobservable factors – which reside in the error term - affecting the diversification choice are not assumed to be the same between the two households, while in the causal interpretation the unobservables are kept unchanged. In other words, in the latter case the ceteris paribus condition includes the unobservables.

This difference in interpretation exists if a relevant variable, which is correlated with the included regressors, is omitted from the model. In other words, if there are factors which influence the likelihood of a household having farm certification which are not captured by the variables currently included in the model. In that case, a causal interpretation of the farm certification variable requires a different estimation method. A possible solution is an instrumental variable approach. An instrument is a variable that is correlated with the endogenous regressor (relevant) but can be assumed to be uncorrelated with the model's error term (valid).

Instrument relevance can be tested directly. It requires that the partial correlation between the instrument and farm certification is non-zero, i.e. that the instrument is significant in the instrumental equation. Instrument validity cannot be tested, making the logic explaining the independence of the instrument with the model's error term all the more important.

Under the assumption that the instrument is valid, the endogeneity of certification can be tested by comparing the regular and the instrumental variable estimators for beta (Hausman, 1978). This was done for both the probit and tobit equations using the same auxiliary regression, often referred to as the Durbin – Wu – Hausman test (Verbeek, 2012). The auxiliary regression was estimated explaining farm certification by all the regressors from equation (1) and (2), X_i and Z_i, as well as the instruments:

$$C_i = \pi_0 + \pi_{1i} I_i + \pi_{2i} X_i + \pi_{3i} Z_i + \nu_i \qquad (3)$$

The saved residuals from this regression, ûi, were added to the models of interest, giving equations of the following form:

$$d_{ij} = \beta_{0j} + \beta_{1j}C_i + \beta_{2j}X_i + \beta_{3j}Z_i + \omega_j \hat{u}_i + \varepsilon_{ij} \quad (4)$$

$$y_{ij} = \beta_{0j} + \beta_{1j}C_i + \beta_{2j}X_i + \beta_{3j}Z_i + \omega_j \hat{u}_i + \varepsilon_{ij} \quad (5)$$

If $\omega_j = 0$, C_i is exogenous. This was formally tested with standard t-tests on ω_j in regressions (4) and (5).

Equation 1 was estimated with the multivariate probit model developed by Cappellari and Jenkins (2003), using the Geweke-Hajivassiliou-Keane (GHK) smooth recursive conditioning simulator (Börsch-Supan and Hajivassiliou, 1993; Keane, 1994). This model is considered an improvement over econometric models previously used to explain participation decisions, which analysed only one activity at the time. Estimating the equations as a system is more efficient when error terms are correlated. For the same reason, equations 2, 4, and 5 were estimated using a Stata package developed by Roodman (2009) based on the same GHK simulator. Tobit models were used when the dependent variable was income, which is continuous, but zero for a substantial part of the range, as is the case for households not obtaining any income from a particular income source (Amemiya, 1984; Tobin, 1958). Robust standard errors were used for all estimations.

Instruments

If there is evidence that farm certification was determined by factors which were not included – an omitted variable problem – in the vector of control variables, a two-stage instrumental variable approach would be required for causal interpretations of its effect on income diversification. However, such an approach only outperforms simpler one-stage estimations when used instruments are relevant and valid (Wooldridge, 2010). Instrument validity is particularly hard to establish, as it requires a variable – ideally several – which affected farm certification but had no independent effect on income diversification.

The historical pattern of farm certification following security improvements as the army pushed back extra-legal armed groups, was described in section 2.2. Unfortunately, no reliable security data was available at a small enough scale to be used as an instrument. Moreover, no data was collected on the year in which farmers received certification, which made it impossible to demonstrate the relationship quantitatively. There was the further concern that guerrilla activity might have a direct effect on income diversification - by increasing the cost of accessing the market and hence encouraging households to become more self-reliant – making the instrument invalid.

In addition to the pattern of certification following security, the high quality of coffee in the region has led coffee buyers to provide their exporters with strong

incentives to certify all supplying farmers. As described in section 2.2, it is traders who select farmers, not farmers who independently decide that certification might be a good way to obtain higher returns. Exporters selling to Starbucks, the largest buyer in the region, receive financial incentives based on the share of supplying farmers which are certified as well as the threat to be excluded from the supply chain if their effort was considered to fall short (Alvarez et al., 2010; Starbucks, 2014). This pressure to certify is not unique to Starbucks and expected to intensify (Elder et al., 2014; Kolk, 2013). Although farmers are certified as rapidly as possible, this does not mean the share of certified farmers in a region goes from zero to one hundred overnight. Which factors might make farmers more likely to be selected for certification while not affecting income diversification directly? More accurate modelling of this farmer selection process would require trader-level data. In this study, only farm-level data was available, requiring identification of a different type of variables which might affect the selection process independently of the diversification decision.

Two sets of variables were identified which might meet the validity criterion. Regional dummies at the municipal level - whose relation with certification and farm size is shown in Table 5.2, and three variables indicative of the degree of social integration of a household. First, the number of years the household had been living in the village. Second, whether or not the households had been displaced by violence. Third, whether the head of the household considered him- or herself native to the community. The hypothesis was that the better a household's social integration, the more likely they were to be selected by traders for certification, because of increased visibility, familiarity, and trust.

Results and discussion

Instrument relevance and endogeneity tests

Instrument relevance was tested with standard t-tests on the instruments in the first-stage equation, reported in Table 5.5. Columns 1 and 2 show first-stage equations of the multivariate probit and tobit models; columns 3 and 4 show results of the auxiliary regression (equation 3). Results show that of the social integration dummies, only the residency years variable was significant. It had the expected sign: the longer households had been living in the village, the more likely they were to hold farm certification. The municipality dummies were jointly significant. Results show households in La Union and Taminango were significantly more likely, and households in Linares significantly less likely to hold farm certification. Control variables, with the exception of household size, added little explanatory power.

The insignificant instruments were removed from the auxiliary regression in

Table 5.5. Instrument relevance: certification status

Equation:	**(1)**	**(2)**	**(3)**	**(4)**
Instruments				
Residency years	0.014***	0.009**	0.014***	0.014***
	(3.49)	(2.17)	(3.32)	(3.30)
Displaced dummy	0.081	0.090	0.135	
	(0.15)	(0.15)	(0.20)	
Native dummy	-0.085	0.036	-0.038	
	(-0.16)	(0.06)	(-0.06)	
Municipality dummies				
Chachagui	-0.440	-0.320	-0.590	-0.606
	(-0.95)	(-0.78)	(-1.35)	(-1.40)
El Tambo	-0.297	-0.231	-0.270	-0.290
	(-0.77)	(-0.59)	(-0.64)	(-0.69)
La Florida	0.484	0.493	0.480	0.466
	(1.32)	(1.37)	(1.24)	(1.21)
La Union	2.761***	3.697***	2.691***	2.658***
	(6.06)	(6.78)	(5.25)	(5.23)
Linares	-0.946*	-0.726	-1.145**	-1.152**
	(-1.95)	(-1.55)	(-2.30)	(-2.32)
Samaniego	-0.032	0.141	-0.107	-0.078
	(-0.07)	(0.34)	(-0.24)	(-0.18)
Taminango	2.781***	2.493***	2.657***	2.640***
	(5.11)	(4.63)	(4.52)	(4.51)
Constant	-1.429**	-1.338*	-1.368	-1.362***
	(-2.03)	(-1.69)	(-1.62)	(-2.69)
Control variables	yes	yes	yes	yes
Method	mv-probit	mv-tobit	probit	probit

Note: n = 495. Dependent variable is farm certification. Baseline municipality is Buesaco. T-statistics are shown in brackets. Stars denote significance at 10, 5, and 1 percent level.

column 4, since they failed to meet the relevance criterion. The regional dummies were jointly significant. The saved residuals from this regression were included in the equations 4 and 5, whose output is presented in Table 5.6. Under the assumption that the instruments are valid - not correlated with the error terms in the diversification equations – standard t-tests on the coefficients on û show whether farm certification is exogenous. Results indicate that this is not the case for only one of the dependent variables in the multivariate tobit regression: coffee income. It may be considered exogenous for all dependents in the multivariate probit model, and for the other dependents in the multivariate tobit model.

Table 5.6. Durbin - Wu - Hausman endogeneity tests

	Self-employed			Wage	
	Agri	**Coffee**	**Non-agri**	**Agri**	**Non-agri**
			Multivariate probit		
û	-0.2	0.1	0.2	0.1	0.3
	(-0.40)	(0.21)	(0.73)	(0.48)	(1.05)
Other variables	yes***	yes***	yes***	yes***	yes***
			Multivariate tobit		
û	-1202.3	8666.5***	1676.2	-164.2	4405.7
	(-1.59)	(5.78)	(0.59)	(-0.24)	(1.31)
Other variables	yes***	yes***	yes***	yes***	yes***

Note: Excerpt from regression output. Only relevant variables are shown.

These results support the narrative of farm certification described in section 2.2, which suggests that the rapid pace at which farmers are being certified in the wake of security improvements, leaves little space for farmers to self-select into certification or for exporters to be overly selective. However, given the relative weakness of the instruments and their strenuous validity, these test-results should be interpreted with some care. On the other hand, if these results were wrong, the same weakness of the instruments would make correcting the apparent endogeneity equally problematic. Assuming the test results are valid, there is no evidence for endogeneity in the probit regressions and some evidence for endogeneity in the tobit regressions. Keeping in mind the pitfalls, for the probit models only the non-instrumented output will be presented and discussed; for the tobit models both the regular, non-instrumented, and the instrumented (second-stage) output will be presented.

Participation decision

Various model specification tests were performed to determine the estimation method, check for necessary variable transformations and heteroskedasticity. The reported Chi-square test rejects a series of independent probit regressions in favour of the multivariate probit model used, supporting the assertion that activity choices are made jointly rather than separately. The estimated correlation coefficients of the residuals presented in Table 5.7 show that this is due to the relationship between coffee production (equation 2) and wage labour (equation 4 and 5).

Further model checks indicated potential non-linearity in the variables land/labour ratio and land holdings. An expanded model including squared terms

of both variables was rejected in favour of a model containing only the squared term of land holdings; only results of the final model are shown. Comparisons of regular with robust standard errors showed little difference between the two. Robust standard errors are reported.

Table 5.8 shows the factors driving participation in five different categories of income-earning activities: on-farm agricultural production other than coffee, coffee production, self-employment outside of agriculture, and wage labour in and outside of the agricultural sector. Coffee producers with farm certification were more likely to currently receive income from coffee or other on-farm agricultural production. Although the size and significance of the coefficient is largest for obtaining coffee income, this effect should be interpreted with caution: smaller coffee producers are less able to spread renovations - resulting in zero coffee production for at least two years on renovated plots - over time, and therefore are more likely to report zero coffee income.

Households with more education were more likely to participate in non-agricultural activities, either through self-employment or wage labour. Education reduced the likelihood of participation in agricultural wage labour and had no effect on the participation decision into coffee or other on-farm agricultural production. This clearly showed the importance of education in overcoming barriers to entry into higher return activities and reducing the need to engage in the lowest return activity, agricultural wage labour.

The variables land/labour ratio, land holdings, and land holdings square were mean-centered, such that the coefficient on land holdings measures the effect of a one-hectare increase for a family with average land and labour holdings (Wooldridge, 2006). Such a family has a little over 2 hectares of land and an adult equivalent household size of 3.3. At these values, an additional hectare of land decreases participation in agricultural wage labour, signals push diversification (Barrett et al., 2001; Ellis, 2000). Households with land holdings too small to generate sufficient income have to work on other farms, even if obtained wages

Table 5.7. Estimated correlation coefficients of the residuals

Equation:	**(2)**	**(3)**	**(4)**	**(5)**
(1)	-0.055	-0.086	-0.055	0.083
	(-1.19)	(-0.87)	(-0.79)	(1.37)
(2)		0.051	-0.120*	0.127**
		(0.77)	(-1.93)	(2.28)
(3)			0.128	0.107
			(1.63)	(1.07)
(4)				-0.025
				(-0.36)

Table 5.8. Participation in specific income-earning activities

	Self-employed			Wage	
	Agri	**Coffee**	**Non-agri**	**Agri**	**Non-agri**
Farm certification	0.602***	1.292***	-0.063	-0.181	0.158
	(2.60)	(5.04)	(-0.41)	(-1.46)	(1.14)
Education	0.013	0.021	0.062***	-0.064***	0.091***
	(0.43)	(0.86)	(2.75)	(-3.48)	(4.28)
Land-labour ratio	-0.003*	0.010*	-0.000	-0.001	0.007**
	(-1.81)	(1.77)	(-0.19)	(-0.76)	(2.31)
Land holdings	0.174	0.071	0.032	-0.127***	0.036
	(1.54)	(1.10)	(0.62)	(-3.04)	(0.75)
Land holdings sq	-0.013**	-0.007	-0.003	0.007**	0.002
	(-2.00)	(-1.62)	(-1.02)	(2.21)	(0.69)
HH size (AE)	0.158**	-0.055	-0.012	0.357***	0.049
	(2.16)	(-0.98)	(-0.23)	(7.09)	(1.02)
Distance to city	0.211***	0.071	-0.072	0.141***	-0.055
	(2.90)	(1.41)	(-1.49)	(3.93)	(-1.27)
Transfer income	0.000	0.000	-0.000	-0.000**	-0.000
	(1.16)	(1.46)	(-0.60)	(-2.28)	(-0.76)
Constant	0.271	0.762***	-1.587***	-0.537***	-1.915***
	(0.83)	(2.83)	(-5.86)	(-2.79)	(-7.65)
Use % (Table 1)	95%	92%	11%	51%	19%

Note: Multivariate probit model. n = 495. T-statistics are shown in brackets. Stars denote significance at 10, 5, and 1 percent level.

are low.

Household size was positively related to engagement in agricultural production and agricultural wage labour. All households in our sample owned land, such that it is likely that excess labour which could not find off-farm agricultural employment was absorbed into on-farm agricultural production (Sadoulet et al., 1998). The importance of household size in explaining participation in on- and off-farm agricultural activities might indicate a shortage of better rewarded income-earning opportunities, including a shortage of opportunities in agricultural wage labour.

Households living further away from cities are more likely to be active in agriculture, both on- and off-farm, evidenced by the significant coefficient on distance to city. Such households are also less likely to own their own business in sectors other than agriculture. However, the distance a households lives from a city does not affect the likelihood it participates in non-agricultural wage labour, which is counter to findings of Lanjouw et al. (2001) for Tanzania and Smith et

al. (2001) for Uganda, who found a negative relationship. It appears that travel distance did not deter those who were offered the opportunity to accept non-agricultural employment.

Households receiving transfer income were less likely to participate in agricultural wage labour, strengthening the image that it is an income source of last resort.

Overall, model results corroborate findings from elsewhere in the livelihood literature, which supported the model specification. All variables added explanatory power, shown by their significance for at least one category of the dependent variable. Model results were strongest for agricultural wage labour, which was in large part due to the equal distribution of this activity between households: 51% of households engaged in this activity. The model was far less powerful for activities with more unbalanced participation rates, as shown by the Use % shown in the bottom row of Table 8, simply because there was less variance to explain.

Generated income

Using the amount of income generated by each activity instead of just the binary participation decision gave the model more explanatory power particularly for those income sources which were used by nearly all households, like coffee production. Model specification tests suggested possible non-linearities in the variables land holdings and distance to cities. Likelihood ratio tests suggested the superiority of the model containing only the squared term of land holdings, which was confirmed by Akaike and Bayesian Information Criteria. Thus, the final multivariate Tobit model (Table 5.9) contained the same variables as the multivariate probit model presented in the previous section.

Having farm certification was related negatively to income obtained from agriculture, positively to coffee income. It had the largest effect on coffee income, which it increased by over 6 million Colombian Pesos (COP), roughly equivalent to USD 3,500. This effect was substantial, given that average household income was 11 million COP, and supportive of earlier findings (Méndez et al., 2010; Ruben and Fort, 2012; Valkila, 2009). However, this was only the effect on income from coffee. Having farm certification reduced income from on-farm agricultural production labour, likely indicating re-allocation of resources away from this activity. Such re-allocation of labour is driven by substitution and income effects. As total household labour is fixed, increased returns to one activity cause households to substitute labour away from other activities. Households are most likely to substitute labour away from activities with low return or which are considered less 'satisfying' (Chayanov, 1966). Income effects lead to an increased consumption of leisure, reducing overall hours worked. The negative impact on income from on-farm agricultural production showed that substitution effects dominated income effects. Coffee income appeared to increase not only through higher prices and

Table 5.9. Income obtained from specific income-earning activities

	Self-employed			Wage	
	Agri	**Coffee**	**Non-agri**	**Agri**	**Non-agri**
Farm certification	-1072.9***	6260.9***	-470.5	-261.5	1104.5
	(-2.84)	(8.00)	(-0.33)	(-0.77)	(0.65)
Education	38.9	417.5***	709.3***	-128.4**	1471.5***
	(0.72)	(3.76)	(3.54)	(-2.53)	(5.79)
Land-labour ratio	-2.4	8.8	6.3	-4.5	21.8
	(-0.42)	(0.75)	(0.35)	(-0.61)	(1.17)
Land holdings	1082.1***	2254.9***	328.2	-239.2*	802.1
	(8.23)	(8.34)	(0.67)	(-1.90)	(1.46)
Land holdings sq	-45.4***	-133.1***	-26.1	15.9*	4.7
	(-4.54)	(-6.43)	(-0.56)	(1.73)	(0.12)
HH size (AE)	-48.1	-185.8	-121.7	1125.5***	623.8
	(-0.36)	(-0.68)	(-0.25)	(9.36)	(1.11)
Distance to city	-296.2***	-670.1***	-602.8	187.4*	-555.4
	(-2.73)	(-2.99)	(-1.44)	(1.94)	(-1.15)
Transfer income	0.0	0.0	0.0	0.0	-0.0
	(1.24)	(0.66)	(0.01)	(0.66)	(-0.54)
Constant	1162.3*	-3722.0***	-15833.2***	-2706.5***	-27161.4***
	(1.90)	(-2.96)	(-5.40)	(-4.74)	(-7.39)

Note: Multivariate tobit model with incomes ('000 COP) as dependent variables. model. n = 495. T-statistics are shown in brackets. Stars denote significance at 10, 5, and 1 percent level.

yields in coffee production, but also through increased allocation of land and labour to its production, reducing the income from sources competing for these resources. This relationship was tested more formally in section 5.4 below.

Education increased income from coffee and non-agricultural activities; it was related negatively to income from agricultural wage labour. Its contribution to income was particularly large in activities outside of agriculture. The negative effect of education on agricultural wage labour is most likely due to the substitution of labour away from this activity. As education is measured by the number of years of education of the most educated household member, this negative relationship implied that households with more highly educated family members relied less on agricultural wage labour, an effect likely strengthened by the perceived low social status associated with the activity.

Asset endowments - land/labour ratio, land holdings, household size - all had the expected effects. Land holdings increased agricultural income and coffee income but reduced income from agricultural wage labour. Households with more labour earned more from agricultural wage labour. Travel distance to the

nearest city reduced agricultural and coffee income, although the latter effect was not significant. Distance reduced the profitability of agricultural production by increasing transport costs. Transfer incomes had no effect on incomes from other sources.

Coefficients on variables other than farm certification are comparable between the one-stage and the two-stage models. Comparing the results of the multivariate tobit models with those of the probit model in section 5.2 shows a stark increase in the number of significant variables, particularly in the equations on income from agricultural and coffee production. The signs on the explanatory variables are as expected and in line with the literature.

To alleviate concerns about endogeneity of the farm certification variable, a two-stage multivariate tobit model was run (Table 5.10). The most substantial change with the one-stage model, shown in Table 5.9, is a large increase in the contribution of farm certification to coffee income, which increased from 6 million COP in the one-stage model to 10 million COP in the two-stage model. This is a tremendous increase, given that average total household income in the sample was 11 million

Table 5.10. Income obtained from specific income-earning activities (instrumented)

	Self-employed			Wage	
	Agri	**Coffee**	**Non-agri**	**Agri**	**Non-agri**
Farm certification	-1435.1***	10384.0***	580.5	-503.0	1886.2
	(-3.18)	(11.43)	(0.33)	(-0.96)	(0.88)
Education	40.7	400.1***	705.4***	-127.8**	1464.7***
	(0.75)	(3.52)	(3.52)	(-2.52)	(5.77)
Land-labour ratio	-2.5	9.5	6.6	-4.6	21.9
	(-0.43)	(0.79)	(0.36)	(-0.62)	(1.17)
Land holdings	1105.4***	1981.0***	269.4	-220.0*	749.3
	(8.34)	(7.12)	(0.54)	(-1.71)	(1.35)
Land holdings sq	-46.5***	-119.8***	-23.1	15.0	7.1
	(-4.63)	(-5.66)	(-0.50)	(1.62)	(0.18)
HH size (AE)	-42.1	-251.6	-148.2	1128.7***	606.5
	(-0.32)	(-0.90)	(-0.30)	(9.38)	(1.08)
Distance to city	-306.9***	-544.0**	-552.5	180.6*	-533.8
	(-2.82)	(-2.37)	(-1.31)	(1.86)	(-1.10)
Transfer income	0.0	0.0	0.0	0.0	-0.0
	(1.21)	(0.81)	(0.03)	(0.62)	(-0.52)
Constant	1296.7**	-5251.9***	-16259.8***	-2614.1***	-27375.1***
	(2.10)	(-4.04)	(-5.42)	(-4.43)	(-7.32)

Note: Two-stage multivariate tobit model with incomes ('000 COP) as dependent variables. model. n = 495. T-statistics are shown in brackets. Stars denote significance at 10, 5, and 1 percent level.

COP. Some caution is warranted in interpreting these values. The reliability of the two-stage model hinges on the quality of the instruments. Although the instruments were tested for relevance (Table 5.6), no such tests exist for validity. If there is an endogeneity problem and instruments had effects on income other than through farm certification, making them invalid, the coefficients in both the one-stage and two-stage model are biased (Hahn and Hausman, 2005). If there is no endogeneity problem, the one-stage model outperforms the two-stage model (Wooldridge, 2010).

Total income effects and impact pathways

Results from the multivariate probit and tobit models indicated farm certification changed the composition of income, increasing coffee income at the cost of other sources of income. However, from these results it was not clear whether overall income increased or decreased as a result of farm certification. That is the topic of

Table 5.11. Total income effects and income pathways

	Total income	**Coffee income**	**Price**	**Yield**	**Land under coffee**
Farm certification	0.37***	2.09***	1.48***	1.26***	0.51***
	(4.68)	(10.19)	(6.79)	(7.17)	(8.74)
Education	0.06***	0.08***	0.03	0.05**	0.01
	(5.54)	(2.90)	(1.00)	(2.50)	(1.56)
Land-labour ratio	0.00	0.00	0.00	0.00	0.00
	(0.91)	(1.19)	(0.90)	(1.11)	(1.24)
Land holdings	0.26***	0.37***	0.10*	0.07	0.27***
	(9.25)	(5.98)	(1.68)	(1.39)	(10.65)
Land holdings sq	-0.01***	-0.02***	-0.01	-0.01*	-0.01***
	(-6.59)	(-4.27)	(-1.61)	(-1.72)	(-6.26)
HH size (AE)	0.06**	-0.12*	-0.08	-0.08	-0.02
	(2.01)	(-1.68)	(-1.07)	(-1.35)	(-1.26)
Distance to city	-0.11***	-0.02	0.11**	-0.03	-0.03
	(-5.03)	(-0.43)	(2.06)	(-0.68)	(-1.57)
Transfer income	7.63***	5.20***	7.00***	4.75***	-1.05***
	(62.02)	(15.53)	(19.80)	(16.63)	(-10.67)
Constant	1296.7**	-5251.9***	-16259.8***	-2614.1***	-27375.1***
	(2.10)	(-4.04)	(-5.42)	(-4.43)	(-7.32)
Adj. R-sq	0.329	0.256	0.094	0.116	0.407

Note: n = 495. Dependent variables in logs. T-statistics are shown in brackets. Stars denote significance at 10, 5, and 1 percent level.

this section. In addition, potential pathways causing the increased coffee income were investigated following the approach applied by Jones and Gibbon (2011) but breaking down their productivity variable into separate variables for yield and land. This approach was implemented by separately regressing price, yield, and land on the full set of explanatory variables. By taking logs, results could be interpreted additively.[9] All regressions were run using Ordinary Least Squares (OLS) with robust standard errors.

Results are shown in Table 5.11. Farm certification increased overall income with 37%. It increased coffee income far more substantially, more than tripling income from this source. Price, yield, and increasing the share of land planted with coffee all contribute to this income increase. Farm certification had a 1.48, 1.26 and 0.51 log point effect on price, yield, and land under coffee, giving a combined effect of 3.25, exceeding the aggregate estimate of its effect on coffee income, 2.09. Price differentials were the most important cause of income differentials between certified and non-certified farmers, accounting for almost half of the income effect. Yield and increased use of land for coffee production accounted for 39% and 16% of the income increase.

Implications

Specialty coffee is rapidly growing in importance and its impact on smallholder coffee producers is tangible. Of the coffee producers interviewed in Nariño, over half had access to these specialty markets through farm certification. Although coffee was their most important income-generating activity, on average it contributed less than fifty per cent to household income. For poorer households, the income contribution of coffee was even less. These figures show the importance of considering the portfolio of income-generating activities when analysing the impact of access to specialty coffee markets, rather than focusing on coffee income alone. To analyse this impact, multivariate estimation techniques were employed, which explicitly allow activity choices to be taken jointly, rather than independently. Specification tests confirmed the validity of this approach.

Results showed a large and significant effect of access to specialty markets on the agricultural activities coffee producers engage in as well as the amount of income derived from them. This relationship between activities indicates competition for household resources between coffee production and other activities. Farm certification increased coffee income at the expense of on-farm agricultural production. No evidence was found of farm certification encouraging diversification into activities other than coffee production, as did for example Barham and Weber (2012). Disaggregating the income effect shows it resulted from a combination of higher prices, higher yields, and increases in land area dedicated to coffee.

Increased allocation of household resources to coffee production resulting from farm certification indicates increasing farm-level specialisation. Such specialisation at producer level fits the established pattern of agricultural development and economic growth (Timmer, 1998), a pattern which also appears to hold for agricultural export-led growth like coffee production (Sanjuán-López and Dawson, 2010). Farm-level specialisation allows for the development of more effective market institutions, bringing down costs and increasing profitability within the local value chain. A factor found to constrain further specialisation is access to land. Current land holdings are insufficient to survive from coffee production alone (Forero Álvarez and Furio, 2010). However, there are few opportunities to buy additional land due to rigidities in the land market. In order for the land required for expansion to become available, opportunities need to be generated in off-farm activities, particularly in non-agricultural wage labour (Winters et al., 2010).

A factor strongly affecting participation in and income from non-agricultural activities is education. Education's positive effect on incomes from almost all sources might be explained by it improving access to information (Pingali et al., 2005). Furthermore, it has a positive effect by reducing dependency on agricultural wage labour, the least remunerated income-generating activity.

An increased dependency on coffee production is not without risks. Coffee prices are nefariously volatile since the International Coffee Agreement was dissolved (Daviron and Ponte, 2005), negatively affecting especially farmers and small-scale traders (Gilbert, 1996). This was made painfully clear by the severity of the so-called coffee crisis (Daviron and Ponte, 2005; ICO, 2003). Coffee, especially coffee arabica, is highly susceptible to contagious disease, the most well-known of which is coffee leaf rust (CLR) (McCook, 2006), which struck Colombia and Central America in the growing season of 2008-11 with disastrous consequences (Cristancho et al., 2012; ICO, 2013). Combating the disease requires chemicals which smallholders can ill afford, as money is needed for food (Morris et al., 2013). Moreover, CLR often kills coffee plants, requiring re-planting. New plants are frequently financed with borrowed money and take three to five years to become productive, creating financial shortfalls lasting many years. Smallholder farmers specialised in and thus heavily dependent on coffee production are unlikely to overcome such shocks without access to adequate insurance and credit facilities.

Certifications have become an integral part of agricultural value chains and their importance is likely to increase in the foreseeable future. The big question for policy makers is how certifications can help to attain rural development objectives. Potential positive impacts of certification include increased value creation at farm level and, through specialisation, to regional value chain development with associated employment generation. Because there are fixed costs to certification, there is a benefit, however small, to scale. To enable farm sizes to grow, it must

be possible to acquire land, which is currently problematic at least partly due to difficulties in obtaining land titles. Moreover, for farmers to convert existing land holdings into coffee production, financing is required, which would also benefit from registered land titles. Credit is important to mitigate risks associated with increased specialization, as are insurance and education. Credit and insurance allow farmers to cope with the risks inherent in increased dependence on coffee production, and credit and education allow households to access higher return activities. In order for certification to truly contribute to rural development, a narrow focus on coffee-related activities alone is unlikely to realize sustainable growth.

Conclusion

The trend towards on-farm certification as a pre-condition for access into specialty coffee markets is having a profound impact on the livelihoods of smallholder coffee producers. We showed that in the study region, there is a significant relationship between farm certification and the participation in and income from several agricultural and non-agricultural income-generating activities. To test this relationship, we used recently developed econometric models which consider these decisions to be taken jointly, rather than independently. Model tests corroborated this approach, confirming that activity choices should not be considered in separation, as was common in earlier research on livelihood activities.

Smallholder coffee producers were found to depend on far more than coffee production alone, on average obtaining less than half their income from coffee. This demonstrates the importance of going beyond prices and yields when analysing the effect of access to specialty markets. Farm certification was found to have a significant relationship with the type of activities households engaged in as well as the amount of income that was derived from them. Certified households allocated more resources to coffee production, reducing their engagement in other agricultural production. Consequently, coffee income increased by more than five times as much as total income. Further specialisation in coffee production was limited by small land holdings, while low education levels were identified as the most important barriers to diversification into non-agricultural activities.

Results of the present study show higher dependency on coffee for coffee producers with access to specialty markets through farm-level certification. Our focus in this paper was access to specialty markets, without considering the relationship between certification and actual sales on the specialty market, an area which warrants further investigation, especially given the trend towards over-certification. Another area requiring further study is the long-term effect of increased dependency on coffee production. Events such as the 2001-03 coffee

price crisis and the 2008-11 coffee leaf rust epidemic show the risk of overly relying on coffee, especially for small-scale producers with little or no financial buffer. The question to what extent the more relational nature of specialty coffee value chains protects producers against such risks remains open.

Acknowledgements

This study was supported by Catholic Relief Services (CRS) and the International Centre for Tropical Agriculture (CIAT) of the Consultative Group on International Agricultural Research (CGIAR). We thank Martha del Rio (CIAT), Stephania Carmona (CIAT), Andres Montenegro (CRS), and Michael Sheridan (CRS) for their comments and support. Furthermore, we would like to thank participants in the 137th Seminar of the European Association of Agricultural Economics for their helpful suggestions.

Appendix A: Power calculations

Sample size was calculated using the following formula:

$$n = \frac{N * Z_{\alpha}^{2} * p * (1-p)}{d^{2} * (N-1) + Z_{\alpha}^{2} * p * (1-p)}$$

where values were chosen based on the information in Table 5.12. The beneficiary population size is given by all beneficiaries in the development project; the population size of the control group is based on the total number of coffee farmers in the study region based on information from the provincial departments of agriculture. In total, 228 beneficiaries and 282 control group farmers were selected for interviewing. According to our calculations, the control group farmers by themselves would have been sufficient to be representative of coffee farmers in each study region. The intervention group farmers should be considered additional observations. Comparisons between both groups to check for sufficient randomization of the attribution of the project showed no observable differences between beneficiaries and control group farmers.

Table 5.12. Assumptions power calculation

Beneficiaries	
Confidence level	0.95
Expected proportion (p)	0.5
Relative precision	0.05
Absolute precision	0.069
Zα at 95% confidence	1.96
Population size (N)	1597
Control	
Confidence level	0.95
Expected proportion (p)	0.5
Relative precision	0.05
Absolute precision	0.065
Zα at 95% confidence	1.96
Population size (N)	11217

Appendix B: List of used variables

Income

Divided into income from self-employment (agriculture, coffee, non-agriculture) and wage (agriculture, non-agriculture). Data was obtained using recall for the twelve individual months preceding the interview date. Reported incomes are the sum over these twelve months.

Farm certification

Dummy variable equal to one when the producer held any farm-level certification providing access to higher value markets such as Nespresso or Starbucks.

Education

Table 5.13. Descriptive statistics

Variable name	Unit	n	mean	SD
Farm certification	dummy	510	0.52	0.50
Education	years	510	8.00	3.75
Land-labour ratio	ratio	510	4.00	32.32
Land holdings	hectares	510	2.05	2.43
Land holdings sq	hectares	510	5.88	31.51
HH size (AE)	adult equivalents	510	3.29	1.50
Distance to city	cost-distance	495	2.03	1.71
Transfer income	COP	510	602805	2511772
Coffee price	COP	510	6359	2481
Coffee yield	kg/hectare	509	686	745
Land under coffee	hectares	510	0.98	0.86

Note: Note: COP = Colombian Peso. 1,798.94 COP = 1 USD (31 May 2012)..

The number of years of education of the household member with most years of education.

Land / labour ratio

Total land holdings divided by family size in adult equivalents (see below).

Land holdings

The total number of hectares the producer has access to through ownership, rental or use rights, including fallow land.

Land holdings squared

The square of land holdings.

HH size (AE)

Household size measured in adult equivalents following the method of Deere and de Janvry (1981), giving weights of one to family members aged 17-59, and lower weights to other family members.

Distance to city

Measured as cost-distance to nearest city of at least 10.000 inhabitants. Cost distance functions took into account road quality and elevation parameters.

Transfer income

Yearly income obtained from public and private transfers.

Coffee price

The price farmers receive per kilogram of coffee.

Coffee yield

The per hectare yield of coffee, measured as total coffee production divided by the land area planted with coffee.

Land under coffee

The land area planted with coffee.

References

Alvarez, G., Pilbeam, C., Wilding, R., 2010. Nestlé Nespresso AAA sustainable quality program: An investigation into the governance dynamis in a multi-stakeholder supply chain network. Supply Chain Management: An International Journal 15, 165-182.

Amemiya, T., 1984. Tobit models: A survey. Journal of Econometrics 24, 3-61.

Ávila, A.F., Escobar, E., Tovar, C.T., 2014. Departamento de Nariño: Tercera Monografía. Fundación Paz & Reconciliación.

Bacon, C., 2005. Confronting the Coffee Crisis: Can Fair Trade, Organic, and Specialty Coffees Reduce Small-Scale Farmer Vulnerability in Northern Nicaragua? World Development 33, 497-511.

Bacon, C.M., Ernesto Méndez, V., Gómez, M.E.F., Stuart, D., Flores, S.R.D., 2008. Are Sustainable Coffee Certifications Enough to Secure Farmer Livelihoods? The Millenium Development Goals and Nicaragua's Fair Trade Cooperatives. Globalizations 5, 259-274.

Bacon, C.M., Sundstrom, W.A., Flores Gómez, M.E., Ernesto Méndez, V., Santos, R., Goldoftas, B., Dougherty, I., 2014. Explaining the 'hungry farmer paradox': Smallholders and fair trade cooperatives navigate seasonality and change in Nicaragua's corn and coffee markets. Global Environmental Change 25, 133-149.

Barham, B.L., Callenes, M., Gitter, S., Lewis, J., Weber, J., 2011. Fair Trade/Organic Coffee, Rural Livelihoods, and the "Agrarian Question": Southern Mexican Coffee Families in Transition. World Development 39, 134-145.

Barham, B.L., Weber, J.G., 2012. The Economic Sustainability of Certified Coffee: Recent Evidence from Mexico and Peru. World Development 40, 1269-1279.

Barrett, C.B., Reardon, T., Webb, P., 2001. Nonfarm income diversification and household livelihood strategies in rural Africa: concepts, dynamics, and policy implications. Food Policy 26, 315-331.

Börsch-Supan, A., Hajivassiliou, V.A., 1993. Smooth unbiased multivariate probability simulators for maximum likelihood estimation of limited dependent variable models. Journal of Econometrics 58, 347-368.

Cappellari, L., Jenkins, S.P., 2003. Multivariate probit regression using simulated maximum likelihood. The Stata Journal 3, 278-294.

Chayanov, A.V., 1966. The theory of peasant economy. Manchester University Press, Manchester, UK.

Cristancho, M.A., Rozo, Y., Escobar, C., Rivillas, C.A., Gaitán, A.L., 2012. Outbreak of coffee leaf rust (Hemileia vastatrix) in Colombia. New Disease Reports 25, 19.

Daviron, B., Ponte, S., 2005. The coffee paradox: Global markets, commodity trade, and the elusive promise of development. Zed Books, New York.

Daviron, B., Vagneron, I., 2011. From Commoditisation to De-commoditisation … and Back Again: Discussing the Role of Sustainability Standards for Agricultural Products. Development Policy Review 29, 91-113.

Davis, B., Winters, P., Carletto, G., Covarrubias, K., Quiñones, E.J., Zezza, A., Stamoulis, K., Azzarri, C., Digiuseppe, S., 2010. A cross-country comparison of rural income generating activities. World Development 38, 48-63.

de Janvry, A., Fafchamps, M., Sadoulet, E., 1991. Peasant Household Behaviour with Missing Markets: Some Paradoxes Explained. The Economic Journal 101, 1400-1417.

de Janvry, A., McIntosh, C., Sadoulet, E., 2010. Fair trade and free entry: Generating benefits in a disequilibrium market.

de Janvry, A., Sadoulet, E., 2006. Progress in the Modeling of Rural Households'

Behavior under Market Failures, in: de Janvry, A., Kanbur, R. (Eds.), Poverty, Inequality and Development. Springer US, pp. 155-181.

De la Hoz, J.V., 2007. Economía del departamento de Nariño: Ruralidad y aislamento geográfico, Documentos de trabajo sobre economía regional. Banco de la República, Centro de Estudios Económicos Regionales (CEER), Cartagena, Colombia.

Deere, C.D., de Janvry, A., 1981. Demographic and social differentiation among Northern Peruvian peasants. The Journal of Peasant Studies 8, 335-366.

Deininger, K., Olinto, P., 2001. Rural Nonfarm Employment and Income Diversification in Colombia. World Development 29, 455-465.

Dercon, S., Krishnan, P., 1996. Income portfolios in rural Ethiopia and Tanzania: Choices and constraints. The Journal of Development Studies 32, 850-875.

Donovan, J., Poole, N., 2014. Changing asset endowments and smallholder participation in higher value markets: Evidence from certified coffee producers in Nicaragua. Food Policy 44, 1-13.

Echeverri, R., 1999. Empleo e ingreso rurales no agrícolas en Colombia, Seminario Latinoamericano sobre desarrollo del empleo rural no agrícola. BID-FAO-CEPAL-RIMISP, Santiago, Chili.

Elder, S.D., Lister, J., Dauvergne, P., 2014. Big retail and sustainable coffee: A new development studies research agenda. Progress in Development Studies 14, 77-90.

Ellis, F., 1998. Household strategies and rural livelihood diversification. The Journal of Development Studies 35, 1-38.

Ellis, F., 2000. Rural livelihoods and diversity in developing countries. Oxford University Press, Oxford.

Elvidge, C.D., Baugh, K.E., Kihn, E.A., Kroehl, H.W., Davis, E.R., 1997. Mapping city lights with nighttime data from the DMSP Operational Linescan System. Photogrammetric Engineering and Remote Sensing 63, 727-734.

Forero Álvarez, J., Furio, V.J., 2010. Colombian Family Farmers' Adaptations to New Conditions in the World Coffee Market. Latin American Perspectives 37, 93-110.

Garcia, J., 2003. Evolución de la distribución de las fincas cafeteras. Hacia una regionalización de la caficultura colombiana. Federación Nacional de Cafeteros de Colombia, Bogotá D.C.

Gereffi, G., Garcia-Johnson, R., Sasser, E., 2001. The NGO-industrial complex. Foreign Policy, 56-65.

Gilbert, C.L., 1996. International commodity agreements: An obituary. World Development 24, 1-19.

Hahn, J., Hausman, J., 2005. Estimation with Valid and Invalid Instruments.

Annales d'Economie et de Statistique, 25-57.

Hausman, J.A., 1978. Specification tests in econometrics. Econometrica 46, 1251-1271.

ICC, 2010. Employment generated by the coffee sector, International Coffee Council. International Coffee Organisation, London, England.

ICO, 2003. Impact of the coffee crisis on poverty in producing countries. International Coffee Organisation, Cartagena, Colombia.

ICO, 2013. Report on the outbreak of cffee leaf rust in Central America and action plan to combat the pest. International Coffee Organisation.

ICO, ITC, CFC, 2000. The gourmet coffee project: Adding value to green coffee. International Coffee Organisation (ICO), International Trade Centre (ITC), Common Fund for Commodities (CFC), London, Geneva, and Amsterdam.

ITC, 2011. Trends in the trade of certified coffees. International Trade Centre, Geneva.

Jones, S., Gibbon, P., 2011. Developing agricultural markets in sub-Saharan Africa: Organic cocoa in rural Uganda. The Journal of Development Studies 47, 1595-1618.

Keane, M.P., 1994. A computationally practical simulation estimator for panel data. Econometrica: Journal of the Econometric Society, 95-116.

Kolk, A., 2013. Mainstreaming sustainable coffee. Sustainable Development 21, 324-337.

Lanjouw, P., Quizon, J., Sparrow, R., 2001. Non-agricultural earnings in peri-urban areas of Tanzania: evidence from household survey data. Food Policy 26, 385-403.

Lanjouw, P., Shariff, A., 2002. Rural non-farm employment in India: Access, income and poverty impact, NCAER Working Paper Series No. 81. NCAER, New Delhi.

McCook, S., 2006. Global rust belt: Hemileia vastatrix and the ecological integration of world coffee production since 1850. Journal of Global History 1, 177-195.

Méndez, V.E., Bacon, C.M., Olson, M., Petchers, S., Herrador, D., Carranza, C., Trujillo, L., Guadarrama-Zugasti, C., Cordón, A., Mendoza, A., 2010. Effects of Fair Trade and organic certifications on small-scale coffee farmer households in Central America and Mexico. Renewable Agriculture and Food Systems 25, 236-251.

Morris, K.S., Méndez, V.E., Olson, M.B., 2013. 'Los meses flacos': Seasonal food insecurity in a Salvadoran organic coffee farming cooperative. Journal of Peasant Studies 40, 457-480.

Nespresso, 2014. The farmers explain the positive impact of the Nespresso AAA

Program in the Cauca Narino region in Colombia.

Oberthür, T., Läderach, P., Posada, H., Fisher, M.J., Samper, L.F., Illera, J., Collet, L., Moreno, E., Alarcón, R., Villegas, A., 2011. Regional relationships between inherent coffee quality and growing environment for denomination of origin labels in Nariño and Cauca, Colombia. Food Policy 36, 783-794.

Pingali, P., Khwaja, Y., Meijer, M., 2005. Commercializing small farms: Reducing transaction costs. ESA Working Paper No. 05-08, FAO.

Ponte, S., 2002. The `Latte Revolution'? Regulation, Markets and Consumption in the Global Coffee Chain. World Development 30, 1099-1122.

Potts, J., van der Meer, J., Daitchman, J., 2010. The state of sustainability initiatives review 2010: Sustainability and transparency. IISD and IIED.

Raynolds, L., 2000. Re-embedding global agriculture: The international organic and fair trade movements. Agriculture and Human Values 17, 297-309.

Raynolds, L., Murray, D., Heller, A., 2007. Regulating sustainability in the coffee sector: A comparative analysis of third-party environmental and social certification initiatives. Agriculture and Human Values 24, 147-163.

Raynolds, L.T., 2009. Mainstreaming Fair Trade Coffee: From Partnership to Traceability. World Development 37, 1083-1093.

Reardon, T., 1997. Using evidence of household income diversification to inform study of the rural nonfarm labor market in Africa. World Development 25, 735-747.

Reardon, T., Berdegué, J., Escobar, G., 2001. Rural Nonfarm Employment and Incomes in Latin America: Overview and Policy Implications. World Development 29, 395-409.

Roodman, D., 2009. Estimating fully observed recursive mixed-process models with cmp, CGD Working Paper. Center for Global Development, Washington D.C., USA.

Ruben, R., Fort, R., 2012. The Impact of Fair Trade Certification for Coffee Farmers in Peru. World Development 40, 570-582.

Ruben, R., Fort, R., Zúñiga-Arias, G., 2009. Measuring the Impact of Fair Trade on Development. Development in Practice 10, 777-788.

Rueda, X., Lambin, E.F., 2013. Linking Globalization to Local Land Uses: How Eco-Consumers and Gourmands are Changing the Colombian Coffee Landscapes. World Development 41, 286-301.

Sadoulet, E., de Janvry, A., Benjamin, C., 1998. Household behavior with imperfect labor markets. Industrial Relations: A Journal of Economy and Society 37, 85-108.

Sanjuán-López, A.I., Dawson, P.J., 2010. Agricultural Exports and Economic Growth in Developing Countries: A Panel Cointegration Approach. Journal of

Agricultural Economics 61, 565-583.

SCAA, 2012. Specialty coffee facts & figures. Specialty Coffee Association of America.

Small, C., 2004. Global Population Distribution and Urban Land Use in Geophysical Parameter Space. Earth Interactions 8, 1-18.

Smith, D.R., Gordon, A., Meadows, K., Zwick, K., 2001. Livelihood diversification in Uganda: patterns and determinants of change across two rural districts. Food Policy 26, 421-435.

Starbucks, 2004. Corporate social responsibility annual report, Seattle, WA.

Starbucks, 2014. C.A.F.E. Practices verifier and inspector operations manual, in: Services, S.G. (Ed.), p. 57.

Timmer, C.P., 1998. The agricultural transformation, in: Eicher, C.K., Staatz, J.M. (Eds.), International agricultural development, Third ed. The Johns Hopkins University Press, Baltimore, Maryland, pp. 113-135.

Tobin, J., 1958. Estimation of relationships for limited dependent variables. Econometrica: Journal of the Econometric Society 26, 24-36.

UNCTAD, 2003. Economic development in Africa. Trade performance and commodity dependence. United Nations Conference on Trade and Development (UNCTAD), Geneva.

Valkila, J., 2009. Fair Trade organic coffee production in Nicaragua — Sustainable development or a poverty trap? Ecological Economics 68, 3018-3025.

Vargas, J.F., Caruso, R., 2014. Conflict, crime, and violence in Colombia. Peace Economics, Peace Science and Public Policy 20, 1-4.

Verbeek, M., 2012. A guide to modern econometrics, Fourth ed. John Wiley & Sons Ltd., Chichester, United Kingdom.

Weber, J.G., 2011. How much more do growers receive for Fair Trade-organic coffee? Food Policy 36, 678-685.

Winters, P., Essam, T., Zezza, A., Davis, B., Carletto, C., 2010. Patterns of Rural Development: A Cross-Country Comparison using Microeconomic Data. Journal of Agricultural Economics 61, 628-651.

Wooldridge, J.M., 2006. Introductory Econometrics: A Modern Approach. Thomson/South-Western, Mason, OH.

Wooldridge, J.M., 2010. Econometric Analysis of Cross Section and Panel Data, 2nd ed. MIT Press Books, Cambridge, Mass.

Notes

1 More information on the website of MARD www.minagricultura.gov.co.

2 More information on the homepage: www.federaciondecafeteros.org.

3 See www.cupofexcellence.org for results of all competitions.

4 The stated interest of companies using farm-level certification in the well-being of the farmers they include, makes such certification a classic example of an inclusive business model.

5 Organic, Fair Trade, and Rainforest Alliance certification were also present in the region, but none of the interviewed farmers with one of these certifications did not also have either C.A.F.E Practices or Nespresso AAA certification, or both. Of the certified farmers, 65% held only a single certification, 29% held two, and 6% three.

6 The data used for this analysis was collected within the framework of the Borderlands Project, implemented by Catholic Relief Services (CRS) (http://coffeelands.crs.org). Power calculations used to determine sample size are presented in appendix A.

7 Coffee plants start bearing fruit only in their third year. Hence, some producers with recently renovated plantations did not report coffee income at the time of the interview.

8 Labour was measured in adult equivalents, using the method of Deere et al (1981).

9 ln(coffee income) = ln(price) + ln(yield) + ln (land).

Chapter 6

General conclusion

Main findings

In this thesis the linking of small-scale farmers in low- and middle-income countries with modern food value chains was investigated. The relevance of this link is rapidly gaining prominence on the world agenda, as a result of the unprecedented population growth and accompanying urbanisation that started during the industrial revolution and is expected to last well into the 21st century. Feeding this urbanized world population will require millions of small farmers to be included in food value chains with extensive food safety and quality controls. There are hundreds of millions of such small farmers worldwide, and most of them live in poverty and food insecurity. If modern value chains would offer higher returns, they might be an important instrument in reducing world poverty. Clearly, if including small-scale farmers was easy we, the world, would already have done it. Yet we have not. Another open question is whether modern value chains offer higher returns and are potential instruments of poverty alleviation. In this thesis both the means of inclusion and its effect were studied. Specifically, how small-scale farmers can be included in modern value chains and how the inclusion affects them.

The different types of relationships between firms and farmers were considered jointly as inclusion mechanisms, defined as organisational structures going beyond arm's-length market transactions governing the economic relationship between a firm and rural households in developing countries, which allows the firm to purchase agricultural products from these households or employ household members in agricultural production or processing. By grouping these mechanisms according to their function – linking small-scale farmers to markets – the analytical focus could be placed on what the different forms have in common and how differences between forms might be explained. There exists a myriad of different inclusion mechanisms, ranging from almost market transactions to almost complete vertical integration. These different mechanisms affect different farmers in different ways, affecting not only outcomes like poverty and food security but also behaviour.

When it comes to inclusion mechanisms, a distinction may be made between 'commercial and inclusive value chains' (Harper et al., 2015) and inclusive business models. Commercial and inclusive value chains are inclusion mechanisms which include the poor, but do not specifically state an interest in improving their well-being. Inclusive business models – the topic of this thesis - do state an interest in improving the well-being of those they include. The thesis is divided into two parts: the first part, consisting of chapters 2 and 3, considers the governance structure of inclusive business models, the second part, consisting of chapters 4 and 5, the measurement of farm-level impact.

In Chapter 2, existing theories on determinants of the type of governance

structure (Williamson, 1996) and the form taken by hybrid arrangements (Ménard, 2004) were applied to three governance structures governing the same transaction: sugar cane sourcing from small-scale farmers in KwaZulu-Natal. It was found that the move from market-based governance to hybrid arrangements could not be explained by only considering Williamson's (1985, 1996) transaction characteristics of frequency, uncertainty, and asset specificity. A more complete understanding would at least require taking production characteristics of the local value chain into account. A comparison of the hybrid arrangements using Ménard's (2004) dimensions of monitoring mechanisms, rent allocation, and enforcement mechanisms provided insight into their inner workings, yet fell short of a complete explanation of their specific form. An aspect not considered by either theory is the role of financial constraints, which was found to be relevant both for explaining the move from market to hybrid governance, as well as for the specific form taken by the hybrids.

The analysis in Chapter 2 showed that the imbalance in investment capacity between firm and farmer caused the ability of the inclusive business model to create value to depend critically on the firm's willingness to investment. This situation was believed to hold more generally for most inclusive business models, and therefore the influence of this dimension was further investigated in Chapter 3. In the context of weak public institutions in which inclusive business models operate, they are expected to rely heavily on safeguards included in the governance structure – private enforcement – rather than public enforcement. Two strands in the organisational economics literature were combined to study this relationship between firm investment and safeguards: transaction cost and positive agency theory. From transaction cost theory, the concepts of asset specificity and appropriability hazard were used to show how the context in which the firm's investment was made influenced its risk exposure. From positive transaction cost theory, the concept of information asymmetry was used to show how production characteristics influenced the degree to which appropriability hazard may be controlled. There was a strong relationship between firm investment and safeguards; more investment meant more safeguards. The type of safeguards appeared to be influenced by production characteristics. In addition, co-investment by farmers was used to signal commitment to the agreement with the firm.

In the second part of the thesis, a contribution was made to improved measurement of farm-level impact. Although the emphasis here was on the impact of value-chain interventions, the results extend to other types of interventions. In Chapter 4, the validity of a common indicator of food security, the Household Dietary Diversity Score (HDDS), was tested. This indicator measures household food access by asking whether certain food groups were consumed in the last 24 hours. The more food groups consumed, the higher is food security. The HDDS was

analysed using the Rasch measurement model, which allows for more accurate evaluation than 'traditional' methods like factor analysis and Cronbach's alpha (Tennant and Conaghan, 2007). Data from Ecuador and Colombia show that the HDDS is not comparable between countries and not necessarily between groups within countries. Moreover, to make the sub-groups within which comparison is possible, detailed information of local dietary patterns is required, which defeats the purpose of using a simple indicator. Even when used within sub-groups, it is still not clear what the indicator measures, as food groups known to matter for food security, such as animal protein, were not or even negatively related to the overall score.[1] Hence, we concluded that the HDDS is not a valid indicator of household food security.

When measuring impact, narrowly focusing on a single indicator or outcome is bad practice not only because the indicator might not be valid or reliable, but also because it risks excluding broader impacts, for example on livelihood choices. In Chapter 5, the effect of farm certification on the composition of income-generating activities of coffee farmers was analysed. Upon certification, coffee income increased substantially as a result of higher prices, higher yields, and a larger share of land dedicated to coffee production. However, overall income increased by less than coffee income. In other words, farmers with certification received more income from coffee production, but less from other activities – in this case income from other crops and agricultural wage labour. These results suggest a re-allocation of labour to coffee production, implying that reaping the benefits of certification requires a willingness of the farmer to spend additional time on coffee production.

Scientific contributions

The main scientific contribution of this thesis is for the first time recognising and studying inclusion mechanisms as a single and heterogeneous group. Earlier work, especially that stemming from de Janvry et al. (1991) and Goetz (1992), recognised that small-scale farmers were excluded from markets because of transaction costs, in particular fixed transaction costs. In this strand of the literature, no differentiation was made between which market these farmers were linked to, nor between types of farmers. Rather, it explained why these farmers were currently excluded from mainstream markets. Specific mechanisms, such as cooperatives and contract farming, were studied as ways to overcome these exclusionary transaction costs. More recently, the rapid rise of supermarkets and consumer certification led to literature studying the impact of such specific mechanisms on the incomes of included small-scale farmers. The question of whether all farmers were able to participate or whether some farmers were excluded was also addressed. However, no research specifically considered these different mechanisms as a group, as a

way to overcome the transaction costs excluding small-scale farmers. This thesis thus contains the first published attempts to analyse the governance structures of inclusion mechanisms as a group.

The governance structures of inclusive business models were jointly analysed through the lens of organisational economics in chapters 2 and 3. Previous research on these models was limited to internal evaluations by donor or implementing organisations and focused on learning outcomes of single or several models implemented within a single project. The lack of rigour and reference to theory in these studies has prevented their publication in the academic literature. Lack of reference to a common framework has also limited the comparability between such studies, as authors describe only those aspects which were considered to be the most relevant for each particular case. As a starting point for building a common framework of inclusion mechanisms, we selected transaction cost theory. This theory has the advantage of placing the analytical focus on the transaction itself – in this case the transaction between small-scale farmer and big business – and thereby framing the problem of governance in terms of common and differentiating characteristics. It also allowed building on previous research on small-scale farmer exclusion from markets. In Chapter 2, the extent to which existing theory could be directly applied to inclusion mechanisms was investigated, showing both which aspects could be explained by existing theory as well as highlighting areas where theory fell short and which requiring further theory-building. In particular, two areas were identified which are not covered by existing theory and require further analysis: the interrelationship between production characteristics of the local value chain and financial constraints of transacting parties.

These areas were explored in more depth[2] in Chapter 3, with particular emphasis on the investment made by the firm. The studied inclusion mechanisms all operated in a weak institutional environment – meaning farm-firm agreements had to rely on private rather than public enforcement – and involved financially constrained small-scale farmers. In these circumstances, the firm was responsible for the largest share of the investment. A strong relationship was found between firm investment, and hence the specific form of the governance structure, and production characteristics – similar to what was found in Chapter 2 – and appropriability hazard: the risk of opportunistic behaviour by farmers. Firm investment in turn had a strong influence on the specific form taken by the governance structure. In general, the higher the investment was, the more safeguards were included in the governance structure. This finding, that financial constraints have such a powerful influence on the specific form of governance structures, has implications for transaction cost theory beyond inclusion mechanisms. It shows that the identity of the partners to a transaction matters, not only ex-post but also ex-ante.

This thesis has also contributed to the discussion on measuring the impact of

development aid in general, and its impact on small-scale farmers in particular. Impact measurement has become increasingly important in the wake of the formulation of the Millennium Development Goals, the Paris Declaration of Aid Effectiveness and the Accra Agenda for Action. Outcomes such as poverty and food security are frequently measured using quick-to-implement and easy-to-use indicators, which are fast and cheap. Given this prominence given to measurement, and the increasing reliance on indicators, it is surprising how little attention has been paid to the validity of these indicators. In Chapter 4, a contribution has been made to the field of impact measurement by testing the validity of one of the most frequently used indicators of food security, the Household Dietary Diversity Score (HDDS). Even though this indicator has been widely employed since its introduction in 2006, no validity test had thus far been conducted. Our finding that the indicator does not pass the validity test reemphasizes the importance of testing validity before using an indicator, or even worse, making it a mandatory component of the results framework of development programs, as was done by USAID (Swindale and Bilinsky, 2006). The validity of the HDDS was tested using Rasch analysis, a methodology developed in the 1960s for educational testing (Rasch, 1960) which is only recently gaining prominence in other sciences like medicine (Smith et al., 2008). Rasch analysis has an advantage over traditional methods because it does not only test whether an indicator is valid, but also allows exploring why this might be the case. Hence, it provides starting points for how the indicator might be improved. We suggest this type of validity tests to become a required step in the development of new indicators, to ensure indicators measure what they intend to measure, in each given population.

Farm-level impact measurement of certification or other supply chain interventions is often limited to outcomes directly related to the relevant crop. In the case of coffee certification, the most frequently studied outcome variables are coffee price, yield, and farm income. However, this narrow focus on crop-specific metrics ignores interactions between behaviour and outcomes. In Chapter 5, household behaviour is taken into account by looking at livelihood activities employed to generate an income. Rural households are known to depend on a variety of activities to make a living (Davis et al., 2010). An explicit link is made between the livelihood and certification literature by analysing the impact of coffee certification on households' portfolio of income-generating activities. Recently developed multivariate probit and tobit models were used to illuminate how farm certification affects the activities households engage in, and the amount of income derived from them. Results clearly show that the impact of certification goes beyond direct income effects on the crop itself, and this most likely takes place through a re-allocation of labour. This demonstrates that although labour input is notoriously hard to measure directly in family farming systems, it may be effectively measured indirectly by applying revealed-preference approaches such

as the one employed in this chapter.

Limitations

The conceptual framework presented in the introduction contains a call to acknowledge heterogeneity at both the inclusion mechanism as well as at the farm level. A large gap remains: in bringing these two levels of heterogeneity together. In the various chapters of the thesis, heterogeneity was only allowed at one level, never at both levels at the same time. The main reason for this was analytical. To analyse the effect of an inclusion mechanism on farm-level response, it is necessary to first understand the component parts or dimensions of the mechanism that is being evaluated. Only once the relevant dimensions are known can their impact be properly investigated. This is one of the main problems of the current literature on inclusion mechanisms, subdivided into supposedly homogeneous groups. When looking at the literature on contract farming for example,[3] some studies find positive impacts, while others find negative impacts. Explaining these differences in impact requires understanding first of all the heterogeneity within the contract farming arrangement itself, and second of the heterogeneity within included farmers. However, the dimensions which might explain these differences are frequently not included in the studies, making ex-post explanation impossible. Until the relevant dimensions are discovered, simultaneous inclusion of heterogeneity at the inclusion mechanism and farm level is not possible.

In order to build as much as possible on previous work, the disparate strands of literature on contract farming, certification, and cooperatives should be brought together under a single theoretical umbrella. At the very least, a theoretical comparison based on archetypes of each mechanism should yield dimensions to include in evaluation studies that would facilitate future comparisons. Ideally, this body of theory would also yield testable predictions. Instead, in this thesis theoretical development was based on non-categorised inclusion mechanisms. The reason for this is two-fold. First, before forcing archetypes of inclusion mechanisms into a particular theoretical mold, it was necessary to prioritize the identification of the extent to which this mold would make an adequate fit or whether further theory-building was required. This is the approach taken in Chapter 2. Second, by focusing on dimensions that are unique or common between archetypes, the heterogeneity within archetypes would have to be ignored. If the heterogeneity within archetypes turned out to be larger than that between them, the approach would be rendered useless. By instead analysing non-categorised inclusion mechanisms, it was more likely that the most important dimensions would be discovered. The extent to which these dimensions differ between archetypes was beyond the scope of this work.

In Chapter 3, which considers how firm investment and appropriability hazards

affect the governance structure of inclusive business models, a factor identified as important in the second chapter is consciously ignored: the role of the government. In almost all inclusive business models, the government or a civil society organisation plays an important role, either as a facilitator, donor, or regulator. What exact role these actors play and how their acting affects the specific form of the governance structure warrants further analysis.[4] In the chapter this factor was ignored due to space limitations. This decision was deemed acceptable since in each of the models studied, responsibility for the on-going management of the initiative clearly lay primarily with the lead firm. However, it is not unthinkable that the role played by the government affected firm investment or the level of safeguards in such a way that models might have changed position in one of the results tables, albeit not in such a way that it would dramatically change the main conclusions.

Future research

Cooperation between public and private sector on development matters is likely to continue,[5] at least into the near future, yet current knowledge about these public private partnerships in agricultural value chains is limited. To properly assess the advantages and disadvantages of different mechanisms for farmer inclusion in different settings, it is necessary to start considering these inclusion mechanisms as a group. Currently the literature is divided, with each mechanism - be it contract farming, certification, or cooperatives - being almost exclusively confined to its own cluster. A common framework is thus urgently needed. It was shown in this thesis that the approach of transaction cost theory works for inclusive business models; extending it to the full set of inclusion mechanisms should be the next step.

An important dimension of inclusion mechanisms which warrants further research is the role of the public sector. Given the relatively recent emergence of public interest in making value chains more inclusive, there is little knowledge about how public funds could be used most effectively to generate an impact at farm level. Governments and civil society organisations want to know which mechanisms work, to determine whether it is worthwhile to increase funding.[6] However, by getting involved, these actors are changing the form taken by the governance structure - as was shown in chapter two – and hence might affect the sustainability of the model and the way it generates an impact at farm level. More research is needed to determine the scope and role of public sector engagement.

A closely related dimension of inclusion mechanisms that was not explicitly considered in this thesis is which farmers may be reached by which inclusion mechanism. An important factor to consider is the investment required to make a business model profitable. Both the exogenous factors driving the investment,

such as crops and intended end-market, as well as the share of the investment which could be borne by outside donors without affecting incentives should be considered. Certain models might only work when farmers co-invest or require a minimum land size. However, not all models have such requirements. Clear identification of factors driving the specific form taken by inclusion mechanism should yield insight into which models work for whom.

The impact of these inclusion mechanisms on the well-being of their intended beneficiaries should be accurately measured. Fortunately, impact evaluations are becoming increasingly common, although - perhaps because such evaluations are relatively new – they often have a narrow focus on a small set of outcome indicators. This narrow focus not only limits our understanding of how the intervention affects behaviour, but in the case of unreliable indicators might lead to erroneous conclusions. The traditional focus of economic statistics on representativeness is no longer sufficient. Economists need to look to other disciplines, like psychology, to step away from a narrow focus on representativeness and start considering construct validity.

There are many ways in which farmer behaviour might be affected by inclusion mechanisms, far too many to measure. Hence, a common framework should be developed within which the governance structures of inclusion mechanisms can be understood. Although in this thesis a few hesitant steps were taken in the right direction, the road ahead is long, perhaps a thousand miles long. Thankfully, the air is thick with promise; it should be a fascinating journey!

Reflections on scientific method

When it comes to studying inclusion mechanisms and their impact on the behaviour and well-being of small-scale farmers, depth and breadth are not mutually exclusive, quite the contrary. In order to understand how inclusion mechanisms attain their objective of linking small-scale farmers to modern value chains and how these mechanisms affect human behaviour, depth naturally follows from breadth. It is therefore necessary to step away from narrow focus areas. Nothing is gained from categorising inclusion mechanisms into narrowly defined groups without understanding the relationship between them. Categorisation should be based on the dimensions which define the governance structure of inclusion mechanisms, not on names given to observed phenomenons. There is a lot of work left to be done, but the work in this thesis shows that approaching inclusion mechanisms as a single group is possible, and fruitful. Although the variety within this group might be dazzling, ignore it we cannot.

To reach the level of detail required to understand the breadth of the variety within the group of inclusion mechanisms, a heavier reliance on qualitative methods is required than is currently common within the field of agricultural economics.

More attention to the details of the inclusion mechanism should not be limited to studies focusing specifically on the governance structure but extend to impact studies. This is not to say we should do away with the quantitative impact studies which are becoming more common. Rather, agricultural economics should become more open towards mixed methods, including qualitative research in the economists' toolbox. Most impact studies are effectively case studies, with high internal but limited external validity. In that respect, they differ little from case studies conducted using qualitative research methods. However, they differ from qualitative case studies by ignoring the deep contextualisation which well-conducted case-study research requires. This limits their usefulness, as comparing results between different case studies is impossible without such contextualisation.

Within quantitative research, there is an ongoing emphasis on representativeness and statistical causation. The importance of representativeness in the context of impact evaluations is sensible, in that proper attribution of project impact requires participants to be identical in both observable and unobservable characteristics from non-participants. One now well-known way to assure this is randomized allocation. Apart from the limitation that randomization might not always be ethical or feasible, sometimes it also just does not matter: it depends on the objective of the study. None of the data in this thesis adhered to the golden rules of randomized controlled trials. The case studies supporting chapters 2 and 3 were selected to capture variety, so-called purposive sampling. The data supporting chapter 4, on the HDDS, were collected within the context of a development intervention, which, not incidentally, is exactly the way the indicator was designed to be collected. Yet, reviewers and conference audiences alike always questioned representativeness. In any science, going against the stream is discouraged. Perhaps we should make sure the stream is broad enough to start with.

Mixed methods research would also assist in strengthening claims for causation. Any statistician will agree that statistics can never prove causation, whether using time series data - post hoc ergo propter hoc - or especially using cross-sectional data - cum hoc ergo propter hoc. Yet we appear obsessed with instrumental variables. Clearly, when an individual states that she took decision A because of B, that is causation. Of course, researchers need to be careful not to solicit desirable answers. Qualitative data are even more sensitive to such bias than quantitative data, yet this bias plagues both. While qualitative data is often dismissed as subjective, quantitative data receives far less scrutiny. Some studies show that potential bias might be substantial (Beegle et al., 2012a; Beegle et al., 2012b), with underreporting of income and agricultural production perhaps the most well-known examples.

Along a similar vein is the issue of construct validity, which is currently almost completely ignored in agricultural economics.[7] Indicators are extremely convenient ways to cost-effectively capture information on policy-relevant

outcomes such as poverty and food security. However, whether they measure what they intend to measure across different contexts is not always certain. Yet, when scrutinizing the methods and data section of publications in agricultural economics, there usually is far more emphasis on representativeness than on construct validity. Most journals do not even allow space to include surveys, nor require online appendices containing them. The danger of not paying sufficient attention to the data collection and extraction process is that real scientific progress is slowed down. In the worst case, it could lead to situations such as the one currently plaguing social psychology, where empirical results which formed the basis of much theory building could not be reproduced (Yong, 2012). Within development, the stakes are high: people's lives are at stake. When the wrong projects get supported because of an inaccurate indicator, less people are helped than would otherwise be the case. Validity matters.

Mind your step

Like governance structures and farmers, policy makers are a heterogeneous bunch, consisting of local governments, international donor governments and organisations, and local and international civil society. Each actor has its own objectives, be it poverty alleviation, economic growth, domestic food production, or even votes. However, to attain any of these objectives, you need farmers. For simplicity it will therefore be assumed here that all policy makers are interested in the effect of the inclusion on the well-being of farmers and on minimizing the cost of sustainably doing so, sustainable in this context meaning 'without requiring further intervention'. Looking at the projects studied in this thesis, this assumption seems reasonable. All actors grouped here as policy makers at least stated that the well-being of farmers was their foremost concern.

Attaining this objective is far from straightforward. When involved as an active third party, the policy maker has no direct stake in the firm-farm agreement. Its direct involvement is only sustained as long as it is bringing money to the table. Since this period is usually only several years, policy makers need to make sure to affect the governance structure in such a way that the inclusion objective remains viable after their involvement ends. This requires striking a fine balance between short and long-term objectives. The sugar cane case studied in Chapter 2 demonstrates how difficult finding this balance is. Local government was directly involved in setting up the cooperative model. In exchange for their financial support, it required the firm to work with local contractors, which did not yet exist. Training contractors, assisting them in getting up and running, and managing them proved so costly that it endangered the sustainability of the entire model. Tellingly, a few year later, when the firm wanted to include farmers in an area further to the north, they chose a different governance structure, without

government funding and without the requirement of having to work with local contractors.

Apart from the short term-long term trade-off, choices need to be made regarding the type of involvement. Providing direct financing, might not be the best way to attain the inclusion objective. In Chapter 3 it was shown that in two case studies in Mozambique, farmers were asked to make a financial commitment before the firm would make theirs. The emphasis in this financial commitment was not on the financial component, but on demonstrating commitment. If they wanted to, the firms could have made the investment. However, by requiring the farmers to make an investment which is only useful if the agreement with the firm holds, and which becomes sunk otherwise, the farmers credibly demonstrated their commitment to keeping their side of the bargain. Such a signal of commitment is less credible if it is made with free money. Another risk of direct financing is that money is fungible. By financing some part of the inclusion mechanism, money is freed up that can now be used elsewhere. Hence, there is a risk of subsidizing firm profit, regardless of what is financed. Although denoting under which conditions direct financing is beneficial falls outside the scope of this thesis, it is clear that it is never without risk.

As an international donor organisation interested in small-scale farmer inclusion, there are many potential projects to choose from. Therefore, it makes sense to set a target which allows for comparing projects. However, care is warranted. First, there is a risk that you are not measuring what you intend to measure. As was shown in Chapter 4, indicators, however sensible they might sound, might measure different things in different contexts, invalidating direct outcome comparisons between projects. Moreover, it might not be clear what exactly is measured. When using indicators, make sure they are tested for validity and reliability, and always use more than one. Second, it is clear that setting targets shapes incentives. As the old dictum goes: "tell me how you measure me and I will behave accordingly". Different targets elicit different behaviour of agents, and therefore might well result in different governance structures. Measuring and setting targets should not be taken lightly.

In addition to direct involvement, policy makers play an important indirect role in shaping inclusion mechanisms through designing the institutional context. Direct support might not always be the best choice. As was shown in Chapter 5, inclusion mechanisms have several effects, not all of which are intended. Furthermore, heterogeneity within inclusion mechanisms makes it difficult to differentiate between projects. Since not all effects are known ex-ante, prudent policy makers might more effectively attain their objective of small-scale farmer well-being through careful design of the institutional context. We know from Chapter 2 how important the institutional context can be. Policies shape the organisational landscape. In Chapter 3 it was shown that weak institutions cause

firms to rely on private safeguards and controls. Such mechanisms are costly, reducing the overall value created within the inclusion mechanism. Moreover, the power in such private control mechanisms is necessarily skewed towards the firm, increasing its bargaining power vis-à-vis farmers. Building a reliable and accessible justice system would reduce the necessity to rely on private controls, increasing the size pie and the share received by farmers. Although this might be a long-term objective, it is something worth striving for.

Wanted: operating manual for spaceship earth

Inclusion mechanisms are a market place. They provide a place for transactions to be conducted and for strangers to meet. Big business and small farmers sometimes appear like they are living on different planets. One drilled by the efficiency demanded to compete on global markets, the other charged with the responsibility of raising a family under the most adverse of circumstances. They are brought together by global forces outside of their own control. Population growth has pushed businesses to look across borders to supply bulging cities with food, one of the basic requirements for human survival. For farmers, population growth has reduced the size of landholdings, requiring them to generate the same value from ever smaller plots of land. The size of this challenge is evidenced by the number of small farmers living in abject poverty. In 2010, of the 3.1 billion people living in rural areas 60% were considered poor and 35% extremely poor (IFAD, 2010). Given these staggering numbers and the on-going challenge of feeding an ever growing world population, how much may we reasonably expect from inclusion mechanisms?

Feeding the world is nothing new, nor is our concern about being able to do so. Our great ability at producing food is as much cause as consequence of the growth in human population numbers. Moving from hunting and gathering to farming – a process started over 10,000 years ago – made it possible to sustain ever denser human populations. Slowly but surely, farming spread to the furthest corners of the world, making hunting and gathering the exception, rather than the rule. By increasing yields and skill in farming on less favourable lands we have managed to make food production keep up with population growth, even if we did not always expect to be able to do so. Yet, a thing or two have changed since Malthus so famously expressed his concerns in his Essay on the Principles of Population (Malthus, 1798). Most importantly, increasing international trade and the integration of the world economy have made food production a global, rather than local issue, rupturing the link between local yields and population density. It is now possible for a country to sustain a high population without having a productive agricultural sector.

A link which has not been ruptured is that between agricultural productivity and

rural income. The more value is created per hectare, the larger is the number of people able to generate an income from that piece of land, or the higher is their income. In most low and middle income countries there is a substantial gap between potential yield and actual yield. The good news is that the technology required to bring yields to their full potential is available, the only difficulty is getting it to the field. Here poverty is throwing a spanner in the works, creating a chicken-and-egg problem. Farmers need money to buy inputs and knowledge, but to get this money they need to produce more, for which they need inputs and knowledge. Directly providing inputs and training has been tried for decades, but this approach did not manage to break the cycle. A new problem presented itself: absence of a reliable market. Without a reliable market, or so it is now believed, farmers will be reluctant to adopt these new technologies, fearing that their investment might not pay off. In this context, inclusion mechanisms were brought to the fore as the silver bullet. Are they?

Ask any scientist a straightforward question and his or her answer will always be the same: it depends. When it comes to the silver bullet properties of inclusion mechanisms, the answer is not much different. For some farmers, they might improve well-being. By providing a clear market and offering the tools necessary to reach it – as the mechanism is designed specifically for the group of farmers it includes – it provides "incentivised learning". However, it seems unlikely that all small-scale farmers will be included or that all included farmers will benefit equally. Although more research is required to determine the fine details, with some models undoubtedly proving to be more inclusive than others, there are at least two factors limiting inclusion and returns: fixed transaction costs and globalised output prices. Fixed transaction costs – those costs which are independent of the size of the transaction – are what inclusion mechanisms are designed to overcome. However, they will never be able to make such costs disappear completely, meaning there is some benefit to scale, however small this benefit may be. The second, and perhaps most important factor limiting what inclusion mechanisms can achieve are globalised output prices, which put a natural limit on how much value one hectare can generate. In order for a farmer to be able to attain a certain level of income, given optimal yields and market prices, a certain size of landholding is required. Below this threshold income level, crops compete directly with family members for resources. Households with income below this threshold level are considered poor.

In economic theory a model has been suggested to capture this dynamic of poverty traps and persistent poverty (Carter and Barrett, 2006). Although it was not specifically developed for farmers or inclusion mechanisms, it demonstrates the process defining the limits of what market-based solutions can achieve. The idea is simple. Households hold assets which they use to generate income. These assets include land and labour, but also include factors which raise their productivity,

such as tools and skills. The total return to these assets fluctuates, meaning that a certain asset base might generate an income that causes the household to live above the poverty line in some years and below it in others. Households can only save, i.e. enhance their asset base, when they are above the poverty line. When they are below the poverty line, they might have to sell assets to survive. How do inclusion mechanisms fit into this picture? Basically, well-designed inclusion mechanisms are like an exogenous shock, increasing the returns to assets for farmers through providing any combination of inputs, knowledge, and higher or more stable prices. Only when this increase in returns is sufficient to lift farmers above the poverty line, here defined as the threshold required for saving to take place, can the firm be certain the farmer is able to keep his side of the bargain.

What this means in practice is that firms will always prefer contracting with farmers with a larger asset base, which usually means more land. Especially those farmers with the ability to make an investment to credibly signal their commitment to the agreement are interesting partners - as was shown in Chapter 3. Moreover, the presence of fixed transaction costs makes the value generated by the transaction – which needs to be shared by firm and farmer – larger for larger farms. This is not to say that inclusion mechanisms will exclude small-scale farmers, just that firms prefer 'larger' small-scale farmers and that there exists a certain minimum size below which firms will not reach. Clearly, the minimum is flexible and depends as much on the mechanism as on the farmer. Most households that are commonly referred to as farmers are more accurately described as rural households with some land for farming. Worldwide, it is estimated that over 50% of the income of rural households is not generated on-farm (Davis et al., 2010). Of course, there is substantial variation between rural households. Some households might be more interested in inclusion mechanisms that allow them to obtain an income from their land without working it themselves. The mechanism described in the sugar cane case study described in Chapter 2 is one such example. Other households however, might want to be farmers but are lacking the opportunities to practice their profession. In the long run, those households need scale to enable them to generate a respectable income from farming alone.

Increases in farm size have historically gone hand in hand with agricultural and economic development. The role of inclusion mechanisms in this trend is ambiguous. Some models appear to retard upscaling, others seem to enhance it. However, their contribution to the upscaling process is likely to be outweighed by their contribution to the transfer of agricultural technology. By design, inclusion mechanisms have an incentive to maximise the value of on-farm production. This is in the mutual interest of firm and farmer. Transferring technology is how this is most effectively achieved. As mentioned before, it is unlikely that all farmers will be reached by inclusion mechanisms. This is not necessarily a limitation. What matters is not as much how many, but which farmers are reached. Farming

is incredibly difficult. In high income countries only a small percentage of the population makes their living as a farmer, and of those only some are doing so successfully. In reality as well as metaphorically, sowing in fertile soil is an important precondition for high yields. Technology should be provided first to those farmers most likely to adopt it. Inclusive business models have an information advantage when it comes to selecting these farmers, and an incentive to find them. Those farmers who become early adopters can serve as an example to others and create local demand for the inputs and knowledge required to close the yield gap. Through this process, inclusion mechanisms might contribute importantly to agricultural development.

That agricultural production needs to increase in low and middle income countries is beyond question. It needs to increase not only to produce the food needed to feed the growing world population, but also for its positive contribution to economic growth. Our success as a species has always been closely tied to our ability to produce sufficient food. Agricultural development is human development. In addition, over the centuries the locus of the problem has shifted from a regional to a global scale. The biggest problems facing the world today are global problems. To solve these problems, we need to start comprehensively considering the world for what it is: our home, spaceship earth. Solving global problems requires working together as a species. Yet, a large share of the world population is not in a position to positively contribute to tackling global problems. This needs to change. Poverty is an impediment to human progress. In the words of Buckminster Fuller:

"*We are not going to be able to operate our spaceship earth successfully nor for much longer unless we see it as a whole spaceship and our fate as common.*"

References

Beegle, K., Carletto, C., Himelein, K., 2012a. Reliability of recall in agricultural data. Journal of Development Economics 98, 34-41.

Beegle, K., De Weerdt, J., Friedman, J., Gibson, J., 2012b. Methods of household consumption measurement through surveys: Experimental results from Tanzania. Journal of Development Economics 98, 3-18.

Carter, M.R., Barrett, C.B., 2006. The economics of poverty traps and persistent poverty: An asset-based approach. Journal of Development Studies 42, 178-199.

Davis, B., Winters, P., Carletto, G., Covarrubias, K., Quiñones, E.J., Zezza, A., Stamoulis, K., Azzarri, C., Digiuseppe, S., 2010. A cross-country comparison of rural income generating activities. World Development 38, 48-63.

de Janvry, A., Fafchamps, M., Sadoulet, E., 1991. Peasant Household Behaviour with Missing Markets: Some Paradoxes Explained. The Economic Journal 101, 1400-1417.

EU development days, 2015. Cecilia Malström, European Commissioner for Trade, published interview.

Goetz, S.J., 1992. A Selectivity Model of Household Food Marketing Behavior in Sub-Saharan Africa. American Journal of Agricultural Economics 74, 444-452.

Harper, M., Belt, J., Roy, R., 2015. Commercial and inclusive value chains: Doing good and doing well. Practical action publishing, Bourton on Dunsmore, UK.

IFAD, 2010. Rural poverty report 2011. International Fund for Agricultural Development, Rome, Italy.

Malthus, T.R., 1798. An essay on the principle of population (Printed for J. Johnson, in St. Paul's Church-Yard, London).

Ménard, C., 2004. The Economics of Hybrid Organisations. Journal of Institutional and Theoretical Economics JITE 160, 345-376.

Rasch, G., 1960. Probabilistic models for some intelligence and attainment tests. University of Chicago Press, Chicago.

Smith, A., Rush, R., Fallowfield, L., Velikova, G., Sharpe, M., 2008. Rasch fit statistics and sample size considerations for polytomous data. BMC Medical Research Methodology 8, 33.

Swindale, A., Bilinsky, P., 2006. Household dietary diversity score (HDDS) for measurement of household food access: Indicator guide (v.2), Food and Nutrition Technical Assistance Project. Academy for Educational Development, Washington D.C.

Tennant, A., Conaghan, P.G., 2007. The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? Arthritis Care & Research 57, 1358-1362.

Williamson, O.E., 1985. The economic institutions of capitalism. The Free Press-Macmillan, New York.

Williamson, O.E., 1996. The mechanisms of governance. Oxford University Press, Oxford.

Yong, E., 2012. Replication studies: Bad copy. Nature 485, 298-300.

Notes

1 This is the equivalent of the worst students passing the hardest question in the exam just as frequently as the best students.

2 In memory of Jacques Cousteau (1910-1997).

3 As we are currently doing for a systematic review on the 'Contract Farming for Increasing Income and Food Security of Smallholder Farmers in Low- and Middle Income Countries' (3ie Systematic Review

– SR6.1088).

4 All the world's a stage, and all the men and women merely players (William Shakespeare, As You Like It, 1599.).

5 In an interview on 26 May 2015, the European Commissioner for Trade re-emphasized the shift from 'a relationship based on development cooperation to a relationship based on trade, investment, and increased economic ties' (EU development days, 2015).

6 An example of this is the funding of systematic reviews on farm-level impact of contract farming and certification by AGRA and DFID (www.3ie.org).

7 Probably equally so in general economics, but I do not know that literature well enough to dare make that assertion.

Samenvatting voor de leek

Grote internationaal opererende ondernemingen zijn er gedurende lange tijd in geslaagd direct contact met kleine boeren in ontwikkelingslanden te vermijden. De hoge vaste kosten per zakelijk contact maakt het interessanter om een klein aantal grote handelspartners te hebben in plaats van een groot aantal kleine partners. Dit heeft geleid tot een duaal systeem met grote kapitaalintensieve landbouwbedrijven die produceren voor zwaar gereguleerde moderne toevoerketens aan de ene kant, en kleine arbeidsintensieve familieboerderijen die produceren voor de lokale markt aan de andere kant. Deze situatie is aan het kenteren. Door de snelgroeiende en naar steden verhuizende wereldbevolking is er een sterk toenemende behoefte aan veilig voedsel, en dit vergt gereguleerde en moderne toevoerketens. Meer land in gebruik nemen is geen optie, omdat bijna alle wereldwijd beschikbare landbouwgrond reeds in gebruik is. Productiviteitsstijging is hierdoor het enige alternatief. De landbouwgrond met de meeste potentie voor productiviteitsstijging is in handen van kleine boeren in lage- en middeninkomenslanden. Het ontbreekt de meeste van deze boeren echter aan de financiële middelen om deze productiviteitsstijging onafhankelijk te realiseren. Integratie in moderne toevoerketens zou een middel kunnen zijn om deze Gordiaanse knoop door te hakken en hun inkomenssituatie te verbeteren. De vraag is hoe.

De transactie van een kleine boer met de volgende partij in de toevoerketen noemt men een integratie mechanisme. In de eerste helft van dit proefschrift ligt de nadruk op zogenaamde 'inclusive business models', een subgroep van integratie mechanismen. Tien modellen uit Burundi, Ethiopië, Kenia, Mozambique en Zuid-Afrika die verschillen in hun organisatiestructuur werden geanalyseerd door middel van transactiekosten theorie. In hoofdstuk twee werd getoetst in hoeverre de transactiekarakteristieken die werden voorgesteld door Williamson en de dimensies van hybride organisaties van Ménard de exacte organisatiestructuur van een casus in Zuid-Afrika konden voorspellen. Alhoewel deze theorieën als vertrekpunt zouden kunnen dienen voor de studie van 'inclusive business models' tonen de resultaten aan dat er nog heel wat werk verzet moet worden voor ze hun organisatiestructuur – hun 'persoonlijkheid' – exact kunnen duiden. Het gebrek aan financiële middelen van kleine boeren lijkt een belangrijke rol te spelen in het verklaren van deze 'persoonlijkheid'.

Welke rol dit gebrek aan financiële middelen speelt bij het integreren van kleine boeren in moderne toevoerketens werd onderzocht in hoofdstuk drie. Om precies te zijn werd er onderzocht wat het gevolg was van een combinatie van investeringen door het aankopende bedrijf en het risico dat de boer met deze investering aan de haal gaat op de waarborgen die onderdeel uitmaken van de organisatiestructuur. Alle bestudeerde casussen opereerden in landen met zwakke overheden, waarin

niet op de lokale rechtspraak vertrouwd kon worden om eventuele conflicten tussen handelspartners op een juiste manier op te lossen. Hierdoor ontstaat een situatie waarin het aankopende bedrijf niet alleen het gros van de investering maakt, maar ook op zichzelf is aangewezen om eventuele problemen voorkomend uit die investering op te lossen. Onze resultaten tonen dat bedrijven significante investeringen alleen dan maken, als ze de hieruit voortvloeiende risico's afdoende intern kunnen afdekken. In situaties waarin het risico oncontroleerbaar lijkt wordt de investering tot een minimum beperkt.

Het tweede deel van dit proefschrift richtte zich op het meten van de impact van dit type integratiemechanismen op het welzijn van boerengezinnen in ontwikkelingslanden. Omdat kleine boeren vaak in armoede leven, hebben de publieke sector en sociaal begane ondernemingen interesse in het achterhalen van het armoedereductiepotentieel van specifieke integratiemechanismen. Jammer genoeg is het meten van impact verre van eenvoudig. Het vereist zorgvuldig en kritisch denken om te achterhalen hoe wat te meten. Gerelateerd hieraan werd in hoofdstuk vier een veelvuldig gebruikte indicator voor voedselzekerheid, de Household Dietary Diversity Score, geanalyseerd door middel van data van ruim duizend boerenhuishoudens in Colombia en Ecuador. Een specifiek soort analyse die nog niet eerder op een soortgelijk probleem werd toegepast (Rasch analyse) werd gebruikt om te bepalen of de vragen waar de indicator uit bestaat hetzelfde concept lijken te meten. De resultaten tonen dat dit niet het geval is. Nauwkeurigere inspectie liet zien dat de prestaties van de indicator verbeterd konden worden door lokale eetgewoontes in acht te nemen en bepaalde vragen die weinig of de verkeerde relatie hadden met het onderliggende concept achterwege te laten. Zelfs als dit gedaan zou worden is het echter nog maar zeer de vraag of de indicator een goede maatstaf voor voedselzekerheid kan bieden.

Alhoewel indicatoren een populaire methode zijn om impact te meten, bestaat er een risico dat ze niet gevoelig zijn voor specifieke gedragsveranderingen als gevolg van integratie in moderne toevoerketens. Dat dit risico reëel is werd aangetoond in hoofdstuk 5. In dit hoofdstuk werd de impact van certificering en de daaruit voortvloeiende toegang tot de markt voor hoogwaardige koffie op het inkomen van een groep van vijfhonderd kleinschalige koffieboeren in het zuiden van Colombia onderzocht. Het principe van 'kostwinning' waar onze analyse op gestaafd werd, neemt alle activiteiten mee die het boerenhuishouden onderneemt om in haar levensonderhoud te voorzien. De resultaten tonen dat gecertificeerde boeren een groter deel van hun inkomen uit koffie halen. Dit komt door hogere productiviteit, hogere prijzen, en door meer land met koffie te beplanten. Echter, hun totale inkomen neemt met minder toe dan hun koffie inkomen. Dit wijst er op dat de tijd en moeite die het kost om certificaten te behalen en te behouden ten koste gaan van andere rendabele activiteiten. Dit soort veranderingen in gedragspatronen dient in acht te worden genomen bij het meten van de impact

van certificering.

Dit proefschrift bevat drie duidelijke bijdragen aan de huidige stand van de wetenschap. Ten eerste, door specifiek nadruk te leggen op de gemeenschappelijke eigenschappen van integratie-mechanismen in plaats van ze als uniek en losstaand te beschouwen, kon de manier waarop deze mechanismen het mogelijk maken handel te drijven met kleinschalige boerenbedrijven uitgebreid worden onderzocht. Dit werd toegepast in hoofdstuk twee, waarin werd aangetoond dat transactiekosten theorie een goed punt van vertrek vormt voor het doen van het soort onderzoek beschreven in dit proefschrif. Echter, bepaalde eigenschappen van integratie mechanismen worden vooralsnog onvoldoende belicht. Met name de rol die het gebrek aan financiële middelen aan de kant van de kleinschalige boerenbedrijven speelt in het vormen van de organisatiestructuur verdient verdere aandacht. In hoofdstuk drie werd een eerste stap in die richting gezet. Samen vormen deze hoofdstukken het eerste wetenschappelijk gepubliceerde werk rond zogenaamde 'inclusive business models'.

Twee verdere wetenschappelijke bijdragen werden geleverd aan het meten van impact op het welzijn van het boerengezin. Ten eerste werd een voor dit onderzoeksveld nieuwe methodologie toegepast in hoofdstuk vier om het functioneren van een veelgebruikte indicator voor voedselzekerheid onder de loep te leggen. Deze methodologie, de eerder genoemde Rasch analyse, staat het toe om de bijdrage van individuele vragen aan het functioneren van de indicator als geheel te bestuderen, iets wat niet mogelijk is met traditionele methodes. Ten tweede werd, in hoofdstuk vijf, het belang van het in acht nemen van gedragsveranderingen op huishoudniveau als gevolg van integratie in toevoerketens aangetoond. Het gros van het onderzoek naar de impact van certificering richt zich op gemakkelijk waar te nemen factoren, zoals prijzen en productiviteit. Een toename van inkomen uit koffie wordt dan al snel geïnterpreteerd als een toename in welzijn. Echter, het onderzoek in dit proefschrift toont aan dat de impact van certificering verder reikt dan alleen het inkomen uit koffie en laat zien hoe een recent ontwikkelde methode gebruikt kan worden om dit zichtbaar te maken.

Deze bevindingen hebben drie duidelijke implicaties voor beleidsmakers. Ten eerste, dat de oneerlijke machtsverhouding tussen kleinschalige boerenbedrijven en grote internationaal opererende ondernemingen voor een deel veroorzaakt wordt door zwakke lokale overheden. Zolang bedrijven niet kunnen vertrouwen op lokale rechtssystemen, is een ongelijke machtsverhouding wellicht de prijs die betaald dient te worden om investeringen in landbouwontwikkeling mogelijk te maken. Ten tweede, dat mogelijkheden voor de publieke sector om dergelijke ongelijkheden weg te nemen beperkt zijn. Als het risico groot is dat een boer met een door het bedrijf gemaakte investering aan de haal gaat, moeten kleinschalige boerenbedrijven op een geloofwaardige manier hun betrokkenheid kunnen aantonen – wat mogelijk is door een investering in de relatie te maken. Een dergelijke

investering toont alleen dan daadwerkelijke betrokkenheid als deze gemaakt is met eigen middelen, niet wanneer deze gemaakt is met gedoneerd geld. Ten derde, dat wanneer beleidsmakers een keuze moeten maken tussen verschillende integratie mechanismen, de manier waarop het mechanisme geëvalueerd wordt zorgvuldig moet worden afgewogen. Er dient een keuze gemaakt te worden aangaande zowel de uitkomsten die meegenomen worden in de evaluatie als de manier waarop deze gemeten worden. Als de beslissing genomen wordt om indicatoren te gebruiken, is het van belang alleen geverifieerde indicatoren te gebruiken. Een evaluatie gebaseerd op een enkele indicator of uitkomst kan tot verkeerde conclusies leiden.

Acknowledgements

Doing a PhD is like solo-travelling. The decisions you take and the consequences flowing out of them are your own responsibility. Yet the experiences and memories you take away from the trip are influenced as much by your own decisions as by the people you meet along the way. And I met a lot of people. Here, I would like to take the opportunity to thank them for their companionship.

First and foremost, my thanks go to Marijke D'Haese, my promotor. She was equally influential in starting me off on this journey, as in providing support along the way. When we first met, Marijke had just started her tenure track, which allowed her to apply for BOF-funding. Lucky for me, she was pregnant at the time and trusted me to write the proposal. The deal was: if the proposal were to be accepted, she would hire me to execute it. The proposal was accepted; the rest, as they say, is history. I would like to thank Marijke for giving me the freedom to do my PhD as I saw fit and supporting me regardless of the decisions I took.

All my colleagues are grand, but there are a few who deserve special mention because of the quality office time we spent together. Sam Desiere, who started his PhD on the same day as me. Over the years we have written several papers together, with the brunt of the analysis always falling on his shoulders. The chapter on the HDDS food security indicator is as much his work as it is mine. Annelien Gansemans, my student turned colleague, who appears shy but is as tough as a nail when it comes to field work and I am sure will make a far better researcher than I will ever be. And Lotte Staelens, the boss of the office, who tried her utmost to limit my talking during work time, with mixed success. The breaks we took to revive our concentration were some of the most pleasurable moments of my time at the department, and the discussions we shared often turned out to be at least as useful as the work time we lost. And my travelling colleagues, Danie Jordaan, Sanctus Niragira, and Leander Raes, who spend most of their time in their home countries, but always embellish lunch with the colourful stories they bring back to the office. Thank you for adding sparkle.

Friday is for drinks with colleagues. Lunch is a social event. These golden rules are the foundation for the contact with my colleagues from the department. Thanks to all of you for making work about more than work.

My PhD process required many travels paid from a diverse host of funds, which at Ghent University implies a lot of administration. Given my almost complete lack of administrative skills, the only reason this process was completed successfully time and time again is thanks to the great team at the secretariat. Without their continuous support, my ship would have stranded long ago in the jungle of paperwork. Special thanks go to Annick, my chief financial officer, Sybille, my

mother away from home, and Jozef, who is like the grandfather who knows where all the tools in the department are (in his desk drawer).

The proposal for my PhD was titled "Smallholder farmer integration into modern agrifood supply chains: methodological innovations" and is built around measuring the impact of a contract farming arrangement of onions in Azerbaijan, a joint Unilever and Oxfam project. In April 2012, at the seas of change conference in The Hague - two weeks after starting my PhD - Jan Kees Vis told me the project would be killed. I needed a beer.

This is how I met Mark Lundy, research group leader Linking Farmers to Market at CIAT. He asked me why I needed a beer so badly, to which my response was "I'm a data analyst without data". His reply was: "We've got loads of data". We became friends immediately. He introduced me to Borderlands, a joint project of CIAT and CRS which aims to improve the livelihoods of around 3,000 coffee farmers on both sides of the Colombian-Ecuadorian border. Since January 2013 l was even promoted lead researcher on this project, a mixed blessing at times when things did not go according to plan. Mark was always the great connector and strategist behind the scenes, putting rocks in the river upstream to make sure we had water where we needed it.

For CIAT, the project team consists of Carolina Gonzalez, Alex Buritica Casanova, and Fernando Rodriguez. All of them are great colleagues. Over time, through various meetings, work trips and shared frustrations we have become friends. Carolina is my 'jefa Colombiana'. I would like to thank her not only for giving me so much responsibility and freedom within the project, but also for encouraging me to hang on, even when things were not going according to plan. Alex, apart from having a great last name, is also a gifted and hard worker. There is nobody I have worked with so intensely during the last years. Since I met him, he has gone from chief photocopier to chief data management and the go-to point for his colleagues regarding any statistical queries. All the work from data management to analysis for the chapter on coffee farmers we did together. Fernando is the oil in the machinery of the team, logistically and financially. He is also the most diplomatically gifted, and great at making clear what he thinks without actually saying it. On top of all of this, he is also my coffee dealer. All the Colombian coffee I have been sharing with you over the years was bought from him.

For CRS, the people I have worked most closely with are Andres Montenegro and Paola Benavides from Colombia, and Carlos Novillo and Alex Portilla from Ecuador, and Michael Sheridan, the project leader. Andres is a great person, and although we always had to diplomatically skirmish to defend our interests, we always had and will have plenty of common ground to fall back on. Paola once sneaked me into Samaniego, which I'm sure nobody would have permitted if they had known. We had a great time. Carlos is a special case, we never got along. Yet, from all the people I have worked with in these years, there is nobody who I

have learned so much from. Most importantly, I have learned that to assume that because you are in the same project you share the same goals is naive. The person and the organisation he works for are two separate things. Thank you Carlos, for making that crystal clear. Alex is the star programmer of the team. Working together on programming the tablets for monitoring and evaluation was a great experience. Michael is the person I have had least direct contact with, although he did interview me once for his coffee blog. Our strongest connection is through the outcomes of the project which he has been so vividly promoting through blogs, presentations, and workshops. Thank all of you for hosting me, working with me, and making me feel part of the team.

The other pillar on which this PhD is build is the pair of projects with CDI and Pretoria University, on inclusive business models. Although there are too many people involved to thank each person individually, a few people deserve recognition. First of all, Jan Helder, for graciously introducing me to many of his contacts in South Africa, including Mandla and Wellington. Without him, it is unlikely my field work would have been in Africa. And of course to Monika Sopov and the project on inclusive business models which provided data for half of this thesis. Our short and to-the-point communication might have been a result of our shared KPMG history.

From Pretoria University, I would like to thank Ward Anseeuw, the most amicable project leader anybody could wish for. From him I learned important rules of thumb that allow combining the personal with the professional, such as 'the number of people minus one is the maximum number of bottles of wine you should put on the tab'. My thanks go to Wytske Chamberlain for always being critical on reduction and simplification and for showing me that it is possible to remain 100% Dutch even if you live abroad. And finally Tijo Salwerda for the quality social time, hiking in wildlife parks while discussing the lifestyle of the Mauritian elite. And of course a great many thanks to Wellington, Ingle, and Bosman for granting me access to their work, facilitating and giving all those interviews.

Finally, two individuals who strongly affected my professional development deserve recognition. First of all Claude Ménard for helping me build the theoretical framework on which chapter 3 in this thesis is built. A big name in institutional economics, he is also an extremely warm and welcoming person, who is truly dedicated to advancing theory. It has been a pleasure to work with him. And then of course Giel Ton, who defended his own thesis on 1 December. Giel invited me to join the systematic review on the impact of contract farming, which introduced me to many of the papers cited in this book. Moreover, I'd like to thank him for the many talks about the difficulty of measuring impact and for his inspiring dedication to making the world a better place for smallholder farmers.

Although my thesis would not have been possible without the direct support of

everybody I worked with, the indirect support of those surrounding me allowed me to sustain my effort. First of all, I would like to thank those who embellished my private life in Gent. My housemate since the first hour, Lynn Jacob, has tried her utmost to integrate me into Flemish culture. She taught me local habits and lingo, and encouraged me to learn-by-doing. The Sunday cava, many dinners, parties, and hung-over days on the couch are among my fondest memories of my time in Ghent. My thanks also go to my squash partners, Diederik and Pieter, for providing an outlet for my energy and a great excuse for having a beer without any women present. Although Pieter could not always make it because of his drawing course, he more than made up for it by designing the cover of this book.

And then of course, there are my Dutch friends. The Monday group Ruben, Jeroen, Willem, Pieter, Pieter, Herman, and Romeo I would like to thank for the food, for both body and mind. Intellectual discussion and good food could not wish for a better home. The gentlemen of Proximitas, to wit Daan, Floris, Jan, Joep, Jurek, Lennaert, Manfred, Peter, Rogier, Ron, Rutger, Thales, Thijs, and Wouter I would like to thank for their teachings on gentlemanly behaviour, at all times. The weekends, the parties, the trips to the Philippines and Riga, were all legendary. Douwe Heeringa, another gentleman, I would like to thank for making Ghent his second home. It does not matter how long we have not spoken, it always seems like yesterday. Most of all, I would like to thank all of you for helping me preserve my Dutch identity, in that faraway land across our southern border.

The older I get, and I have gotten a lot older recently, the more I have come to value family. The loss of my grandfather, to which this book is dedicated, was a big blow. I was in Colombia for work when it happened and therefore one of the few people who could not be at his bedside while he died. He was like a second father. I thank my parents, Piet Vellema and Joke de Lange, for always being there for me and for providing the certainty that I have something to fall back on. The same goes for my sister Ditte and brother Jelmer. I highly value the advice of each of you. Moreover, your persistence in asking me when I would finally finish my PhD certainly helped me to keep up the pace and maintain focus. My big family, from both my mother's and my father's side I would like to thank for the laughs, the light conversation, and the mutual understanding that whatever happens we will always support each other.

The many travels which I had to make in the context of my PhD resulted not only in input for my thesis, they also enriched my personal life. The most vivid example of this is my girlfriend, Katherine Gonzalez Valdes, who I met, albeit only for a few minutes, at a fastfood stand in Colombia. I would like to thank her for supporting me, by coming to Belgium but also by absorbing the frustrations I brought home when I felt things were going too slow at work. Her patience has been an inspiration.

Making mistakes is inevitable, and I'm grateful to Lotte Staelens, Eline D'Haene,

Annelien Gansemans, Jelmer Vellema, Piet Vellema, Joep Hoveling and Ruben Mars for taking the time to meticulously check the final version of my thesis for grammatical errors. Indubitably, some errors still remain, but without them it would have been many more. All errors, both those removed and those remaining, are my sole responsibility.

Some people have the gift to be able to be brief. I do not think I am one of those people. I have tried to keep these acknowledgements as short as possible, but as I was writing I kept remembering more and more people to thank. Nevertheless, not everybody could be mentioned. My apologies to those who I felt obliged to ignore for sake of brevity; you are not forgotten. To all of you who I have mentioned and all of you who I have not, I would like to express my heartfelt appreciation for making this PhD a memorable journey!

Never forget: "*The pleasure we derive from journeys is perhaps dependent more on the mindset with which we travel than on the destination we travel to.*"

Alain de Botton, The Art of Travel

List of abbreviations

B2C	Business to consumer
CDI	Centre for development innovation
CFSM	Core food security module
CGIAR	Consultative group for international agricultural research
CIAT	International centre for tropical agriculture
CLR	Coffee leaf rust
CRS	Catholic relief services
DANE	National administrative department of statistics (Colombia)
DIF	Differential item functioning
ELCSA	Latin American and Caribbean food security scale
ESS	Equity share scheme
FANTA	Food and nutrition technical assistance
FAO	Food and agriculture organisation
GHK	Geweke-Hajivassiliou-Keane
HDDS	Household dietary diversity score
HFIAS	Household food insecurity access scale
ICC	Item characteristic curve
ICO	International coffee organisation
IFAD	International fund for agricultural development
IIED	International institute for environment and development
INEC	National institute of statistics and censuses (Ecuador)
IRF	Item response function
NAMC	National agricultural marketing council
OECD	Organisation for economic cooperation and development
OED	Oxford English dictionary
PPI	Progress out of poverty index
RA	Rainforest alliance
RM	Rasch model
RV	Recoverable value
SA	South Africa
SASA	South African sugar association
SASRI	South African sugarcane research institute
THS	Tongaat Hulett sugar
UN	United nations
WFP	World food programme

Curriculum vitae

Wytse Vellema was born in Ferwerderadeel, The Netherlands, on 19 September 1985. He has two siblings, a sister and a brother. After high school, he obtained a BSc in international economics and business (honours, cum laude) from the Rijksuniversiteit Groningen (RUG), a postgraduate diploma in economics from Cambridge University, and an MSc international development with a specialization in development economics (cum laude) from Wageningen University. During his studies, he was an active member of a student association, culminating in a year of being its chairman in 2007-08. He has work experience in corporate finance (KPMG) and second-tier financing of microcredit and farmer cooperatives (Oikocredit). His main interest is the interface between business and small-scale agriculture in developing countries, specifically in finding business models which are mutually profitable and have scaling potential. Wytse is (co-) author of several papers published in academic journals in the field of agricultural and development economics. He presented his work at numerous national and international conferences.

PUBLICATIONS AND BOOKS

Published

Vellema, W., S. Desiere, M. D'Haese. Verifying validity of the Household Dietary Diversity Score: An application of Rasch modelling. Accepted for publication in the March 2016 issue of the Food and Nutrition Bulletin

Vellema, W., D'Haese, M., 2015. Explaining hybrid "personalities" in smallholder sugar cane sourcing. British Food Journal 117, 2547-2563.

Vellema, W., Buritica Casanova, A., Gonzalez, C., D'Haese, M., 2015. The effect of specialty coffee certification on household livelihood strategies and specialisation. Food Policy 57, 13-25.

Desiere, S., Vellema, W., D'Haese, M. "A validity assessment of the Progress out of Poverty Index (PPI)™." Evaluation and program planning 49 (2015): 10-18.

BOOK: Blomne Sopov, M.; Saavedra Gonzalez, Y.R.; Sertse, Y.; Vellema, W.; Verjans, H. 2014. Is Inclusive Business for you? Managing and upscaling an inclusive company : Lessons from the field, Wageningen UR (University & Research Centre). Wageningen.

Vellema, W., Mujawamariya, G., D'Haese, M. (2014). Gum arabic collection in northern Kenya: Unexploited resources, underdeveloped markets. Africa

Focus, 27(1): 69-86.

Vellema, W., Mujawamariya, G., D'Haese, M., Burger, K. (2013). An economic approach to household collection of gum arabic from the wild. International Forestry Review, 15(2): 255-269.

Working papers

Desiere, S., Vellema, W., D'Haese, M. Beyond averages: using survey data for informed policy making. Submitted.

Vellema, W., Ménard, C., D'Haese, M. Inclusive business models: How investment and appropriability hazard shape governance structure.

BOOK: Chamberlain, W., Anseeuw, W., Vellema, W. Inclusive Business Models in South African agriculture. Under revision by publisher.

PRESENTATIONS AND WORKSHOPS (by date)

15-16 juli 2015. 146th EAAE Seminar, Ghania (GR). Inclusive business models: how investment and appropriability hazard shape governance structure.

Vellema, W., Ménard, C., D'Haese, M. 8-10 juni 2015. EAAE PhD workshop, Rome (IT). Inclusive business models: how investment and appropriability hazard shape governance structure. Vellema, W., Ménard, C., D'Haese, M.

29 april 2015. BAAE PhD workshop, Brussels (BE). Inclusive Business Models: Beyond contract farming

8-13 september 2014. World coffee science conference ASIC2014, Armenia (CO). Man versus Machine: Measuring cup quality with NIR. Vellema, W., Tolassa, K., D'Haese, M., Boeckx, P.

26-29 augustus 2014. EAAE conference, Ljubljana (SI). Verifying validity of the Household Dietary Diversity Score: An application of Rasch modelling. Vellema, W., S. Desiere, M. D'Haese.

19-23 mei 2014. Workshop ESNIE2014, European School of New Institutional Economics, Cargese (FR). Presented theoretical approach to analysis of inclusive business models.

30 april 2014. BAAE PhD workshop, Brussels (BE). Inclusive Business Models: Beyond contract farming. Vellema, W., D'Haese, M.

25-26 april 2014. GlobalFood symposium, Göttingen (DE). Inclusive Business Models: Beyond contract farming. Vellema, W., D'Haese, M.

4 april 2014. Stakeholder workshop "Inclusive development projects", Pretoria

(SA). Presented case studies on Bosman Family Vineyards and Tongaat Hulett Sugar.

3 October 2013. Opening workshop "Inclusive development projects", Pretoria (SA). Presented methodology for case studies developed in cooperation with CDI, Wageningen UR. All 18 case studies in the project were conducted using this method.

11-13 september 2013. 138th EAAE seminar "Pro poor innovations in food supply chains", Gent (BE). Income diversification and specialty coffee, a trade-off or win-win? Vellema, W., A. Buritica Casanova, C. Gonzalez, M. D'Haese.

7 december 2012. GAPSYM 6, Gent (BE). On collecting gum Arabic in the wild: An economic approach. Vellema, W., Mujawamariya, G., D'Haese, M., Burger, CJP.